Pertanika Journal of

SCIENCE & TECHNOLOGY

JST

VOL. 26 (4) OCT. 2018

PERTANIKA JOURNALS

A scientific journal published by Universiti Putra Malaysia Press

Journal of Science & Technology

About the Journal

Overview

Pertanika Journal of Science & Technology (JST) is the official journal of Universiti Putra Malaysia published by UPM Press. It is an open-access online scientific journal which is free of charge. It publishes the scientific outputs. It neither accepts nor commissions third party content.

Recognized internationally as the leading peer-reviewed interdisciplinary journal devoted to the publication of original papers, it serves as a forum for practical approaches to improving quality in issues pertaining to science and engineering and its related fields.

JST is a **quarterly** (January, April, July and October) periodical that considers for publication original articles as per its scope. The journal publishes in **English** and it is open to authors around the world regardless of the nationality.

The Journal is available world-wide.

Aims and scope

Pertanika Journal of Science and Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: bioinformatics, bioscience, biotechnology and bio-molecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management, and related fields of study.

History

Pertanika was founded in 1978. A decision was made in 1992 to streamline Pertanika into three journals as Journal of Tropical Agricultural Science, Journal of Science & Technology, and Journal of Social Sciences & Humanities to meet the need for specialised journals in areas of study aligned with the interdisciplinary strengths of the university.

After almost 25 years, as an interdisciplinary Journal of Science & Technology, the revamped journal now focuses on research in science and engineering and its related fields.

Goal of *Pertanika*

Our goal is to bring the highest quality research to the widest possible audience.

Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions are guaranteed to receive a decision within 14 weeks. The elapsed time from submission to publication for the articles averages 5-6 months.

Abstracting and indexing of *Pertanika*

Pertanika is almost 40 years old; this accumulated knowledge has resulted in Pertanika JST being abstracted and indexed in SCOPUS (Elsevier), Thomson (ISI) Web of Knowledge [BIOSIS & CAB Abstracts], EBSCO & EBSCOhost, DOAJ, ERA, Cabell's Directories, Google Scholar, MyAIS, ISC & Rubriq (Journal Guide).

Future vision

We are continuously improving access to our journal archives, content, and research services. We have the drive to realise exciting new horizons that will benefit not only the academic community, but society itself.

Citing journal articles

The abbreviation for Pertanika Journal of Science & Technology is *Pertanika J. Sci. Technol.*

Publication policy

Pertanika policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. It prohibits as well publication of any manuscript that has already been published either in whole or substantial part elsewhere. It also does not permit publication of manuscript that has been published in full in Proceedings.

Code of Ethics

The Pertanika Journals and Universiti Putra Malaysia takes seriously the responsibility of all of its journal publications to reflect the highest in publication ethics. Thus all journals and journal editors are expected to abide by the Journal's codes of ethics. Refer to Pertanika's **Code of Ethics** for full details, or visit the Journal's web link at http://www.pertanika.upm.edu.my/code_of_ethics.php

International Standard Serial Number (ISSN)

An ISSN is an 8-digit code used to identify periodicals such as journals of all kinds and on all media–print and electronic. All Pertanika journals have ISSN as well as an e-ISSN.

Journal of Science & Technology: ISSN 0128-7680 (*Print*); ISSN 2231-8526 (*Online*).

Lag time

A decision on acceptance or rejection of a manuscript is reached in 3 to 4 months (average 14 weeks). The elapsed time from submission to publication for the articles averages 5-6 months.

Authorship

Authors are not permitted to add or remove any names from the authorship provided at the time of initial submission without the consent of the Journal's Chief Executive Editor.

Manuscript preparation

Refer to Pertanika's **INSTRUCTIONS TO AUTHORS** at the back of this journal.

Most scientific papers are prepared according to a format called IMRAD. The term represents the first letters of the words **I**ntroduction, **M**aterials and Methods, **R**esults, **A**nd, **D**iscussion. IMRAD is simply a more 'defined' version of the "IBC" [Introduction, Body, Conclusion] format used for all academic writing. IMRAD indicates a pattern or format rather than a complete list of headings or components of research papers; the missing parts of a paper are: *Title, Authors, Keywords, Abstract, Conclusions*, and *References*. Additionally, some papers include Acknowledgments and Appendices.

The *Introduction* explains the scope and objective of the study in the light of current knowledge on the subject; the *Materials and Methods* describes how the study was conducted; the *Results* section reports what was found in the study; and the *Discussion* section explains meaning and significance of the results and provides suggestions for future directions of research. The manuscript must be prepared according to the Journal's **INSTRUCTIONS TO AUTHORS**.

Editorial process

Authors are notified with an acknowledgement containing a *Manuscript ID* on receipt of a manuscript, and upon the editorial decision regarding publication.

Journal of Science & Technology
Journal of Science & Technology

Pertanika follows a **double-blind peer-review** process. Manuscripts deemed suitable for publication are usually sent to reviewers. Authors are encouraged to suggest names of at least three potential reviewers at the time of submission of their manuscript to Pertanika, but the editors will make the final choice. The editors are not, however, bound by these suggestions.

Notification of the editorial decision is usually provided within ten to fourteen weeks from the receipt of manuscript. Publication of solicited manuscripts is not guaranteed. In most cases, manuscripts are accepted conditionally, pending an author's revision of the material.

As articles are double-blind reviewed, material that might identify authorship of the paper should be placed only on page 2 as described in the first-4 page format in Pertanika's **INSTRUCTIONS TO AUTHORS** given at the back of this journal.

The Journal's peer-review

In the peer-review process, three referees independently evaluate the scientific quality of the submitted manuscripts.

Peer reviewers are experts chosen by journal editors to provide written assessment of the **strengths** and **weaknesses** of written research, with the aim of improving the reporting of research and identifying the most appropriate and highest quality material for the journal.

Operating and review process

What happens to a manuscript once it is submitted to *Pertanika*? Typically, there are seven steps to the editorial review process:

1. The Journal's chief executive editor and the editorial board examine the paper to determine whether it is appropriate for the journal and should be reviewed. If not appropriate, the manuscript is rejected outright and the author is informed.

2. The chief executive editor sends the article-identifying information having been removed, to three reviewers. Typically, one of these is from the Journal's editorial board. Others are specialists in the subject matter represented by the article. The chief executive editor asks them to complete the review in three weeks.

   Comments to authors are about the appropriateness and adequacy of the theoretical or conceptual framework, literature review, method, results and discussion, and conclusions. Reviewers often include suggestions for strengthening of the manuscript. Comments to the editor are in the nature of the significance of the work and its potential contribution to the literature.

3. The chief executive editor, in consultation with the editor-in-chief, examines the reviews and decides whether to reject the manuscript, invite the author(s) to revise and resubmit the manuscript, or seek additional reviews. Final acceptance or rejection rests with the Edito-in-Chief, who reserves the right to refuse any material for publication. In rare instances, the manuscript is accepted with almost no revision. Almost without exception, reviewers' comments (to the author) are forwarded to the author. If a revision is indicated, the editor provides guidelines for attending to the reviewers' suggestions and perhaps additional advice about revising the manuscript.

4. The authors decide whether and how to address the reviewers' comments and criticisms and the editor's concerns. The authors return a revised version of the paper to the chief executive editor along with specific information describing how they have answered' the concerns of the reviewers and the editor, usually in a tabular form. The author(s) may also submit a rebuttal if there is a need especially when the author disagrees with certain comments provided by reviewer(s).

5. The chief executive editor sends the revised paper out for re-review. Typically, at least one of the original reviewers will be asked to examine the article.

6. When the reviewers have completed their work, the chief executive editor in consultation with the editorial board and the editor-in-chief examine their comments and decide whether the paper is ready to be published, needs another round of revisions, or should be rejected.

7. If the decision is to accept, an acceptance letter is sent to all the author(s), the paper is sent to the Press. The article should appear in print in approximately three months.

   The Publisher ensures that the paper adheres to the correct style (in-text citations, the reference list, and tables are typical areas of concern, clarity, and grammar). The authors are asked to respond to any minor queries by the Publisher. Following these corrections, page proofs are mailed to the corresponding authors for their final approval. At this point, **only essential changes are accepted**. Finally, the article appears in the pages of the Journal and is posted on-line.

Pertanika Journal of

SCIENCE & TECHNOLOGY

Vol. 26 (4) Oct. 2018

PERTANIKA JOURNALS

A scientific journal published by Universiti Putra Malaysia Press

JST
Journal of Science & Technology
AN INTERNATIONAL PEER-REVIEWED JOURNAL

EDITORIAL OFFICE

JOURNAL DIVISION
Office of the Deputy Vice Chancellor (R&I)
1st Floor, IDEA Tower II
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 Serdang, Selangor Malaysia.
Gen Enq.: +603 8947 1622 | 1616
E-mail: executive_editor.pertanika@upm.my
URL: www.journals-jd.upm.edu.my

PUBLISHER
UPM Press
Universiti Putra Malaysia
43400 UPM, Serdang, Selangor, Malaysia.
Tel: +603 8946 8855, 8946 8854
Fax: +603 8941 6172
E-mail: penerbit@putra.upm.edu.my
URL: http://penerbit.upm.edu.my

ABSTRACTING AND INDEXING OF PERTANIKA JOURNALS

Pertanika is almost **40 years old**; this accumulated knowledge has resulted in the journals being abstracted and indexed in **SCOPUS** (Elsevier), Clarivate Analytics [*formerly known as Thomson (ISI)*] **Web of Science™ Core Collection-** Emerging Sources Citation Index (ESCI). Web of Knowledge [BIOSIS & CAB Abstracts], **EBSCO** and EBSCOhost, **DOAJ**, **ERA**, **Google Scholar**, **TIB**, **MyCite**, Islamic World Science Citation Center (ISC), ASEAN Citation Index (ACI), **Cabell's Directories** & Journal Guide.

Pertanika Journal of Science & Technology
Vol. 26 (4) Oct. 2018

Contents

Foreword

Welcome to the Fourth Issue 2018 of the Journal of Science and Technology (JST)!

JST is an open-access journal for studies in Science and Technology published by Universiti Putra Malaysia Press. It is independently owned and managed by the university and run on a non-profit basis for the benefit of the world-wide science community.

This issue contains 30 regular articles. The authors of these articles come from different countries, namely Malaysia, Indonesia, Iran, Iraq, Nigeria, India, Algeria, Turkey and France. India alone contributed 17 articles, the highest number of articles.

Articles submitted in this issue cover various scopes of Science and Technology including applied sciences and technologies, chemical sciences, earth sciences, engineering sciences, environmental sciences, information, computer and communication technologies and mathematical sciences.

Two articles are selected from the most favourable scope in this issue (information, computer and communication technologies). One is an article entitled "Automatic Annotation of Educational Videos for Enhancing Information Retrieval" by *Poornima, N.* and *Saleena, B*., fellow researchers from School of Computing Science and Engineering, VIT Chennai, India. The article discussed the accessibility of educational videos as one of the best measn in imparting knowledge to learners. The study aimed to extract keywords from videos and to find association between the extracted terms. The study had proven the effectiveness of the proposed approach based on the experimentations done on actual videos set. Details of the study is available on page 1571.

The other one selected from the same scope is an article entitled "WBMFC: Efficient and Secure Storage of Genomic Data" by *Sanjeev Kumar, Suneeta Agarwal* and *Ranvijay*, fellow researchers from Department of Computer Science and Engineering, MNNIT Allahabad, India. The article proposed an approach of applying encryption during compression to efficiently and securely store massive genomic data in fasta/multi-fasta file format. The study showed that the proposed algorithm (WBMFC) outperformed the state of the art algorithms in terms of processing time and compression ratio both. Insight of the article is available on the page 1913.

Selected from the scope of mathematical science is an article entitled "Statistical Estimators as an Alternative to Standard Deviation in Weighted Euclidean Distance Cluster Analysis" by *Paul Inuwa Dalat*u and *Habshah Midi*, fellow researchers from Adamawa State University, Nigeria and Universiti Putra Malaysia (UPM), Malaysia, respectively. The

article focused on clustering, one of the major sources of primary data mining tools. The study suggested two methods to overcome the weakness of Standard Weighted Euclidean Distance method used for clustering. It was proven that the proposed methods (namely *Qn* and *Sn* Weighted Euclidean Distance) evidently outperformed the existing methods. The details of the study is available on page 1823.

Selected from the scope of engineering sciences is an article entitled "Experimental Investigations and Exergetic Assessment of 1 kW Solar PV Plant" by *Ranjana Arora* and *Rajesh Arora*, fellow researchers from Amity University Haryana, India. The study aimed to assess the performance/efficiency of the plant in their respective institution. The study applied two exergy techniques based on concepts of thermodynamics and chemical/photonic energy of input solar insolation. The study concluded that SPV systems were direct energy conversion devices which were eco-friendly and environmentally sustainable sources of energy that could yield huge amount of power at minimum operating/maintenance cost. The details of the study is available on page 1881.

We anticipate that you will find the evidence presented in this issue to be intriguing, thought-provoking and useful in reaching new milestones in your own research. Please recommend the journal to your colleagues and students to make this endeavour meaningful.

All the papers published in this edition underwent Pertanika's stringent peer-review process involving a minimum of two reviewers comprising internal as well as external referees. This was to ensure that the quality of the papers justified the high ranking of the journal, which is renowned as a heavily-cited journal not only by authors and researchers in Malaysia but by those in other countries around the world as well.

In the last 12 months, of all the manuscripts processed, 36% were accepted. This seems to be the trend in Pertanika Journals.

We would also like to express our gratitude to all the contributors, namely the authors, reviewers and editors, who have made this issue possible.

JST is currently accepting manuscripts for upcoming issues based on original qualitative or quantitative research that opens new areas of inquiry and investigation.

Chief Executive Editor
Prof. Dato' Dr. Abu Bakar Salleh
executive_editor.pertanika@upm.my

Pertanika J. Sci. & Technol. 26 (4): 1547 - 1570 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Influence of Discrete Fibers and Mesh Elements on the Behaviour of Lime Stabilized Soil

Shenbaga R. Kaniraj[1]* and Y. C. Fung[2]

[1]*Formerly Faculty of Engineering, Universiti Malaysia Sarawak, 94300, Malaysia*
[2]*Song Kau Construction Sdn. Bhd, No. 9, 1st Floor, Lot 16774, Tabuan Tranquility, Jalan Canna, 93350 Kuching, Malaysia*

ABSTRACT

Addition of chemical binders such as lime and cement improves the strength and stiffness of fine grained soils. However, the treated soils exhibit brittle stress-strain behaviour. Inclusion of randomly oriented discrete fibers in the soil-binder mixture changes its brittle behaviour into ductile behaviour. Most synthetic fibers, however, tend to get entangled and cannot be easily separated from one another. Therefore, it is difficult to realize soil-binder-fiber mixtures in which the fibers are distributed uniformly throughout the mass. This issue has been an impediment in the utilization of the positive modification in the behaviours of soils and soil-binder mixtures by the fibers. The present study aims to address the limitations in using fibers as soil reinforcement. Further, it also aims to investigate the use of synthetic mesh or net elements as an alternative type of soil reinforcement. The paper presents the experimental study on a fine grained soil. Lime has been chosen as the binder due to its low cost and the scarcity of fiber reinforced soil studies in which lime has been used as a binder. The main experimental program is a series of unconfined compression tests on samples prepared using untreated soil, soil-reinforcement mixture, soil-lime mixture, and soil-lime-reinforcement mixture. The lime treated samples were cured up to 120 days at laboratory temperature. The results demonstrate the combinational effects of lime and discrete reinforcement elements on the behaviour and mechanical properties of the soil. The performances of the fiber and mesh element reinforcements have also been compared.

Keywords: Fiber reinforcement, fine grained soil, lime stabilization, mesh reinforcement, reinforced soil, unconfined compressive strength

ARTICLE INFO

Article history:
Received: 02 May 2017
Accepted: 27 July 2018
Published: 24 October 2018

E-mail addresses:
kaniraj@hotmail.com (Shenbaga R. Kaniraj)
fungyuetcher@yahoo.com (Y. C. Fung)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Soils exhibit poor mechanical characteristics as they consist of unbound discrete particles. One of the common methods of improving the mechanical behaviour and properties of fine grained soils is chemical stabilization using binders such as lime, cement, and other industrial waste products. Earlier on in the process of stabilization, the chemical stabilizers modify the plasticity characteristics of the soil and improve its workability. Later, in the presence of water in the soil, the chemical binders react and through hydration and pozzolanic reactions form new products that bind the soil particles together which improves the strength, and stiffness of the soil. The stress-strain behaviour of the treated soil, however, becomes brittle and sudden failure occurs at small strain levels.

Fiber soil reinforcement is a technique of mechanical stabilization of soils where the soil is blended with discrete nonreactive short fibers. Various types of natural and synthetic fibers can be used. Compared to the untreated soil, the fiber-soil composite has not only a higher strength but also a more ductile stress-strain behaviour. Addition of both fibers and binders has a synergistic effect on the stress-strain behaviour of soils. First, the strength of the soil is increased. Second, the fibers modify the brittle stress-strain behaviour of the soil stabilized by binders alone into ductile behaviour.

Even though many research studies have established the beneficial effects of fibers in modifying the mechanical behaviour and properties of soils, it has not yet been possible to translate these benefits into practice; the main reason for this being the difficulty in mixing the fibers with the soil to get a soil-fiber or soil- binder-fiber mixture in which the fibers are distributed uniformly throughout the mass of the mixture. Most fibers tend to lump together and cannot be separated into individual fibers easily. New types of soil reinforcing elements that would bring about the same beneficial effects of fibers but at the same time can be mixed with the soil relatively easily are required. The present study focuses on this particular need or problem.

Past Studies on Soil-binder-fiber Mixtures

The study of fiber reinforced soil has attracted the attention of researchers since the 1970s. Legeay et al. (1972) in France and Yang (1972) in U.S.A. reported the earliest studies on the behaviour of reinforced soil. McGown (1978) and Verma and Char (1978) reported the earliest studies on fiber reinforced granular soils. Since then there have been a number of investigations on different aspects of fiber reinforced soil. These studies can be distinguished into two categories as, (a) studies on fiber reinforced coarse grained soils, and (b) studies on fiber reinforced fine grained material.

The behaviours of fiber reinforced coarse grained soils of different sizes ranging from fine sand to coarse sand and silty sand, of different gradations ranging from uniformly graded to well graded sands, and prepared at relative densities (D_r) ranging from 34% to

71%, have been studied by employing different testing techniques such as direct shear test, unconfined compression test, and undrained and drained triaxial shear tests. Despite its origin in the 1970s, research on fiber reinforced granular soil is continuing in the 21st century also (Ahmed et al., 2010; Babu et al., 2008; Consoli et al., 2009, 2010; Diambra et al., 2010; Hamidi & Hooresfand, 2013; Kalumba & Chebet, 2013; Li, 2005; Li & Zornberg, 2013; Park, 2009; Pino & Baudet, 2015; Rao & Nasr, 2012; Sadek et al., 2010; Santoni et al., 2001; Santoni & Webster, 2001; Shao et al., 2014; Yetimoglu & Salbas, 2003).

Similarly, several researchers have investigated fiber reinforced fine grained material (Botero et al., 2015; Butt et al., 2016; Chauhan et al., 2008; Chen et al., 2015; Consoli et al., 2012; Correia et al., 2015; Cristelo et al., 2015; Estabragh et al., 2012; Fatahi et al., 2012; Kaniraj & Havanagi, 2001; Kaniraj & Gayathri, 2003, 2006; Kim et al., 2008; Kumar & Gupta, 2016; Li, 2005; Maheshwari & Solanki, 2009; Mirzababaei et al., 2013; Olgun, 2013; Oliveira et al., 2015; Park & Tan, 2005; Plé & Lê, 2012; Qu & Sun, 2016; Tang et al., 2007; Yi et al., 2015). The fine grained materials investigated include silt, clay, fly ash, and soil-fly ash mixtures. Specimens of fine grained materials were prepared in maximum dry density-optimum moisture content (MDD-OMC) and non MDD-OMC states. As in the case of coarse grained soils, different types of tests such as direct shear test and undrained and drained triaxial shear tests had been carried out. Unlike the sand specimens, the fine grained material specimens can retain their shapes without any lateral support. Therefore, unconfined compression test had been used most commonly and other types of tests, namely tensile test and flexural test, had also been conducted.

In some studies of both coarse grained soils and fine grained materials, only fibers were mixed with the soil. In some other studies, both chemical binders and fibers were used. Research was concentrated on fiber reinforced granular soils till year 2000; studies on fiber reinforced fine grained material gained impetus only in the new millennium. Salient features of the past studies and major conclusions from them are discussed in the following sections.

Types of Fibers Used. An important characteristic of the fibers is their chemical composition which governs their physical, chemical and mechanical properties. A range of natural, synthetic, metallic and other types of fibers have been used in different studies. Table 1 lists the different types of fibers used in each of these four categories of fibers.

Natural fibers are environmentally friendly. From engineering point of view, however, natural fibers suffer some limitations. They tend to absorb water and increase in size. This reduces the moisture content of the wet soil during compaction. They also tend to shrink upon drying which can affect the bond between the fiber and the surrounding soil. Natural fibers degrade with time, particularly in adverse environmental conditions, for example in acidic conditions. Degradability limits the use of the natural fibers to short term

applications. Further, natural fibers are not uniform, vary in their mechanical properties and are flammable. Natural fibers can be treated to improve their unfavorable qualities, but this increases the cost of the fibers. Typical properties of the natural and synthetic fibers used in some recent studies are shown in Tables 2 and 3, respectively. The tensile strength and modulus of natural fibers are lower than those of synthetic fibers.

The major disadvantage of the metallic fibers is their high cost. Because of their high unit weights, relatively more fiber content compared to natural and synthetic fibers is required. Some metals too can corrode in acidic environment.

Synthetic fibers, which do not have the limitations of the natural fibers, have been preferred in majority of studies. Synthetic fibers have uniform characteristics; high melting point, tensile strength, and modulus; and are chemically stable in adverse environmental conditions. Of the different types of synthetic fibers, as shown in Table 1, the most extensively investigated is the polypropylene fiber. Polypropylene fibers are cheaper compared to the other synthetic fibers (Hoover et al., 1982). Polypropylene fibers also have high elongation at break. In the various studies conducted using polypropylene fibers, there was significant variation in the tensile strength and modulus of the fibers - tensile strength varying in the range of 120 – 517 MPa, and tensile modulus in the range of 3000 – 6000 MPa. The variation in the tensile properties of the fibers is bound to influence the results of the studies.

Table 1
Types of fibers used in research studies

Natural (14)	Synthetic (44)	Metallic (13)	Others (10)
Bhabar (1)	Polyamide (nylon) (3)	Aluminum (foil, rod) (3)	Bungi cord (1)
Coir (4)	Polyester (6)	Copper (wire) (3)	Carpet fiber (2)
Human hair (1)	Polyethylene (fiber, mesh, strip) (2)	Galvanized steel (1)	Fiberglass (2)
Oil palm empty fruit bunch (1)	Polypropylene (crimpled fiber, fibrillated fiber, monofilament fiber, mesh, pulp, tape) (30)	Steel (rod, wire) (5)	Fishing wire (1)
Palmyra (3)		Stainless steel (1)	Linen (1)
Reed (3)			Paper (1)
Rubber (1)	Polyvinyl alcohol (1)		Parachute chord (1)
Wheat straw (1)	Polyvinylchloride (1)		Wood (dowel, rod) (1)
	Polyethylene terephthalate (1)		

The number inside the parentheses shows the number of investigations in which that type of fiber had been used.

Table 2
Typical properties of natural fibers

Fiber type	Specific gravity	Tensile strength, MPa	Tensile modulus, MPa	Elongation at break, %	Source
Coir	-	102	2,000	-	Babu et al. (2008)
Coir	0.85	100	2,000	24	Chauhan et al. (2008)
Oil palm empty fruit bunch	1.46	283	-	15	Ahmad et al. (2010)
Oil palm empty fruit bunch (coated)	1.43	306	-	19	
Coir	1.40	60-90	-	30	Anggraini et al. (2015)
Wheat straw	0.1	110 N	-	22	Qu and Sun (2016)
Human hair	1.25-1.4	400	-	-	Butt et al. (2016)

Table 3
Typical properties of synthetic fibers

Fiber type	Specific gravity	Tensile strength, MPa	Tensile modulus, MPa	Elongation at break, %	Source
Polypropylene (monofilament)	0.91	517	3,400	-	Santoni et al. (2001a)
Polypropylene (fibrillated and tape)	0.91	310	4,800	-	
Polyester	1.30	80 – 170	1,450 – 2,500	-	Kaniraj and Havanagi (2001) Kaniraj and Gayathri (2003)
Polypropylene	0.91	320 – 400	3,500 – 3,900	-	Yetimoglu and Salbas (2003)
Polypropylene	0.91	350	3,500	-	Tang et al. (2007)
Polypropylene	0.91	150	3,000	-	Chauhan et al. (2008)
Polypropylene (monofilament)	0.91	120	3,000	80	Consoli et al. (2009, 2010, 2012)
Polypropylene (crimped)	0.91	225	-	160	Diambra et al. (2010)
Polypropylene	0.91	330 – 370	3,500	16 – 20	Tang et al. (2007)
Polypropylene (monofilament)	0.91	400	6,000	-	Hamidi and Hooresfand (2013)
High density polyethylene	0.743	15 – 20	389.7	-	Kalumba and Chebet (2013)
Polypropylene	0.905	250	3,500 – 3,900	-	Correia et al. (2015)
Polypropylene	0.91	120	3,000	80	Chen et al. (2015)
Polypropylene	0.9 to 0.91	-	-	-	Kumar and Gupta (2016); Olgun (2013)

Fiber Reinforcement Parameters

The three important fiber reinforcement parameters are fiber content, fiber length, and, aspect ratio. There has been again significant variation in the values of these parameters in various studies.

Fiber Content

The most common definition of fiber content, f_c, is the ratio of the weight of the fibers to the weight of dry soil or dry soil-binder mixture and expressed in percent. Thus,

$$f_c = \frac{W_f}{W_s} \times 100 \qquad [1]$$

In Eq. 1, W_f = weight of fibers, W_s = weight of dry soil or dry soil-binder mixture. Other less common definitions of fiber content are in terms of volume-volume, area-area, and weight-volume relationships.

Volumetric fiber content, f_{cv}, is defined as,

$$f_{cv} = \frac{V_f}{V_s} \times 100 \qquad [2]$$

In Eq. 2, V_f = volume of fibers, V_s = volume of soil-fiber or soil-binder-fiber mixture.

Fiber content in terms of area, f_{ca}, is useful while considering a particular area of cross-section. It is defined as,

$$f_{ca} = \frac{A_f}{A_s} \times 100 \qquad [3]$$

In Eq. 3, A_f = area of fiber, A_s = total area of soil-fiber or soil-binder-fiber mixture.

Fiber content in terms of weight to volume, f_{cw}, is defined as,

$$f_{cw} = \frac{W_f}{V_s} \times 100 \qquad [4]$$

Fibers usually form a very small component in the soil-fiber matrix. In 82% of the past studies, f_c was ≤ 1%. Hoover et al. (1982) used as small as f_c = 0.02% in their study. Several investigators concluded the optimum f_c as < 1% and as low as 0.3%.

Fiber length, l

While some investigators used fibers of different lengths to study the effect of l on the reinforced soil behaviour, some others had used only fibers of constant length. The length of the fibers used in most of the studies was small. In nearly 40% of the past studies, l ranged between 5 and 15 mm. In two-thirds of the studies, l was ≤ 25 mm and in 87% of

the studies it was ≤ 50 mm. Mirzababaei et al. (2013) used the smallest fiber length of 2 mm in their study. Santoni et al. (2001a) concluded that the optimum value of l was 51 mm.

Aspect ratio

Aspect ratio is the ratio of the fiber length to the diameter or thickness of the fiber. In 49% of the past studies the aspect ratio was ≤ 200 and in 76% of the studies it was ≤ 400. Qu and Sun (2016) used aspect ratios as small as 1.25 in their study. Rao and Nasr (2012) from their study using linen fibers with aspect ratios of 50, 100 and 150, recommended aspect ratio of 100 for use. From the past studies, it can be concluded that fibers of length of up to 50 mm and aspect ratio up to 400 can be used in fiber reinforced soil applications.

Effect of Fibers on Soils

The salient findings of past studies on the effect of fibers on granular soils and fine grained material are discussed separately in the following sections.

Effect of Fibers on Granular Soils

Chemical stabilization of sands is uncommon. Of the 33 studies reported on fiber reinforced granular soils, cement was used as a binder only in five studies. The cement content (C_c) was low (≤ 4%) in three of them (Consoli et al., 1998; Hamidi & Hooresfand, 2013; Park, 2009) and a maximum of 7% and 10% in the other two (Consoli et al., 2009, 2010). The findings, in general, are summarized below. The conclusions from some studies were, however, different.

1. Fibers increased the unconfined compressive strength (UCS). The increase was linear up to a certain value of f_c.
2. The fibers imparted ductility to the sand which manifested through increase of axial strain at failure (ε_f) and reduction in the post peak strength loss.
3. Fibers changed the strain softening nature of stress-strain curves into strain hardening nature.
4. Fibers produced a spongy effect which reduced the MDD and initial secant modulus (E_i).
5. The energy absorption capacity as measured by the area under the stress-strain curve was increased by the fibers.
6. In cemented sands, fibers decreased the brittleness index (I_B) as defined in Eq 5. Decrease in I_B indicates increase in ductility or decrease in brittleness.

$$I_B = \frac{q_p}{q_u} - 1 \qquad [5]$$

In Eq. 5, q_p is the peak or failure stress and q_u is the ultimate or residual stress.

Effect of Fibers on Fine Grained Material

Unlike sands, fine grained soils exhibit plasticity due to the surface characteristics of the particles. Chemical stabilization with binders such as lime and cement with low binder content is therefore common in fine grained soils. The binder content is usually defined in terms of weight-weight relationship as in Eq. 1: the ratio of weight of binder to the weight of dry soil. Some investigators have used the weight-volume relationship, weight of binder to the volume of soil, in their studies. While cement has been used only in 15% of the past studies on granular soils, in the case of studies on fine grained material binders have been used in one-half of the past studies. Cement was the common stabilizer used. Cement content, C_c, varied from 2% to 5% in 45% of the studies, 8% to 10% in another 40% of the studies, and 15% to 20% in the rest 15% of the studies. Lime had been used only in one study. Also, one-fourth of the studies investigated the tensile behaviour of the fine grained material-binder-fiber composite material. The general findings are summarized below.

1. The studies varied in their conclusions on the effect of fibers on the characteristics of fine grained material such as compaction, UCS, rigidity or E_i, ductility, and tensile behaviour. The influence of the fibers was reported variously as significantly beneficial, marginally beneficial, beneficial subject to specific conditions, and even adverse. It is much more difficult to prepare uniform soil-binder-fiber mixtures in fine grained materials than in granular soils. This probably could have influenced the outcome of the different studies.
2. The binders contributed to the improvement of the soil behaviour. The increase in UCS due to both fibers and cement was either more than or nearly equal to the sum of the increase caused by them separately.
3. The variations in the conclusions of different studies indicate that there is a need for more research on fiber reinforced fine grained material and for research using other types of binders such as lime.

Experiments on Lime-Fiber and Lime-Mesh Stabilized Soil

Background of the Study. A constraint to the successful adoption of fiber reinforced soil in practice is the difficulty in realizing soil-fiber mixtures in which the fibers are distributed uniformly throughout the mass. The problem is even more severe in fine grained soils. Preparing even small quantities of uniform soil-fiber mixtures in laboratory studies is difficult and time consuming. Any additional time involved in field operations will also entail additional cost. The fibers should remain untangled, be easily mixable with the soil, and be also not expensive. It may be easier to mix other forms of reinforcing elements such as net or mesh elements with the soil, because for the same reinforcement content the number of mesh elements per unit volume will be less than the number of fibers. The

grid-like two-dimensional structure of the meshes can also contribute to better interlocking between the soil particles. Only a few studies have been reported using mesh elements. McGown (1978) conducted the earliest study using aluminum meshes. Al-Refeai (1991) found the performance of the mesh elements to be superior to that of glass fibers. Santoni et al. (2001a), however, reported that meshes contributed only to a small increase in strength compared to the fibers. But, they used very large size meshes, 51 x 102 mm, in their study. Kim et al. (2008) used waste fishing net with relatively large mesh openings, 22 x 22 mm. Kalumba and Chebet (2013) conducted direct shear tests on sands mixed with high density polyethylene (HDPE) solid strips and strips with perforations of 1 mm and 2 mm diameter on 6 mm wide strips. They reported that the perforations increased the friction angle, φ. Where a binder also was used, cement has been the choice of most investigators. Consoli et al. (2012) used lime as binder. Lime is cheaper than cement. The present study was undertaken keeping in mind all the factors mentioned hereinbefore. Accordingly, fiber and mesh elements were obtained from an inexpensive insect net and lime was used as the binder. The details of the study and the results are explained in the subsequent sections.

MATERIALS AND METHODS

Materials

The three principal materials used were a fine grained soil, lime, and synthetic reinforcement.

Soil Properties

Disturbed soil was collected by the side of Kuching-Kota Samarahan Expressway at coordinates of 1°28'58.2"N 110°24'45.0"E. Laboratory tests conforming to standard test procedures (BS 1377, 1990) were conducted on the soil. The test results are shown in Table 4. Figure 1 shows the grain size distribution of the soil. According to the Unified Soil Classification System, the soil was classified as high plastic silt with sand (MH). As the organic content was < 1%, no additional additive was needed to counter the cationic exchange capacity (Texas Department of Transportation 2005).

Properties of Lime

Fresh calcium oxide (CaO), also called quicklime, manufactured by SIGMA-ALDRICH was used in the study. The pH of the calcium oxide was 12.40 at 25℃. This is within the range of 12.35 to 12.45 recommended by BS 1924 (1990) for lime used in soil stabilization. In the initial consumption of lime test, soil samples were mixed with lime contents, C_l, varying from 5% to 14% and tested for their pH values. Figure 2 shows the results. The pH increased with C_l and reached a maximum value of 12.4 at C_l = 9%. Further increase in C_l did not increase the value of pH. Therefore, C_l = 9% was chosen as the binder content in the study.

Addition of lime (C_l = 9%) influenced the compaction characteristics of the natural soil. It had relatively a minor influence on OMC compared to that on MDD. Lime increased the OMC of the soil from 21% to 22% and decreased the MDD from 1,680 kN/m^3 to 1,560 kN/m^3. The compaction characteristics of the natural soil and the soil mixed with C_l = 9% are shown in Figure 3.

Table 4
Properties of the natural soil

Property	Value
Natural water content, w_n	28.04%
Organic content	0.95%
D_{60}	0.02 mm
D_{30}	0.004 mm
D_{10}	0.0018 mm
Specific gravity, G	2.58
Liquid limit, w_l	50%
Plastic limit, w_p	28.8%
Plasticity index, I_p	21.2%
Shrinkage limit, w_s	3.9%
pH	5.8
Maximum Dry Density, *MDD* (Standard Proctor)	1680 kN/m^3
Optimum moisture content, *OMC* (Standard Proctor)	21%

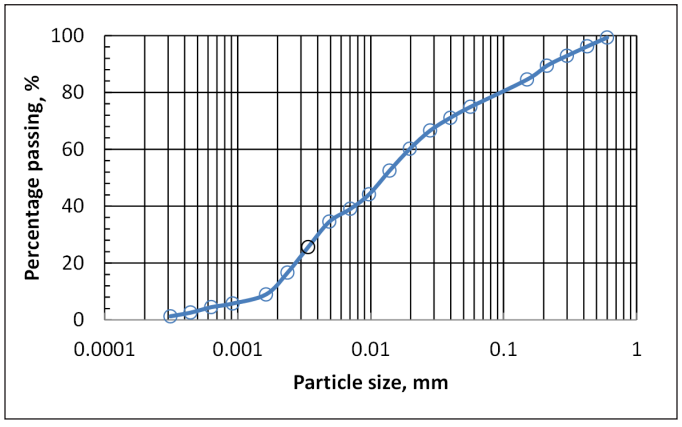


Figure 1. Grain size distribution of soil

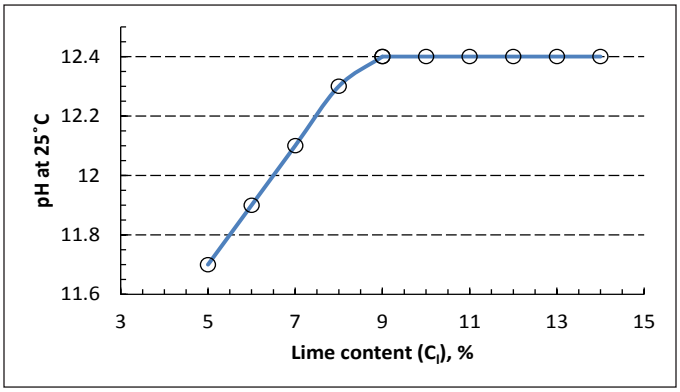


Figure. 2. Results of initial lime consumption test

Fibers and Mesh Elements

Fibers and mesh elements were cut from a green color synthetic insect net having 2 mm x 2 mm diamond shape mesh opening. As the mesh opening was more than the largest soil particle size (≈ 0.6 mm), the mesh would be completely embedded in the soil. The insect nets are usually made of HDPE; they are UV stabilized, and rot resistant. The diameter of the unconfined compression test specimen was 35 mm. Therefore, the lengths of the fibers and meshes were chosen as 7 mm. The mesh was cut in a diamond shape with each side of the diamond as 7 mm long. There were 9 openings in a single mesh element. The diameter or thickness of the fibers and meshes was 1 mm. Figure 4 shows the scanning electron microscope (SEM) images of the fiber and mesh.

Details of the Experimental Study

Experimental Variables. Fiber content and mesh content are designated as f_c and m_c, respectively, in this paper. Both are defined according to the weight-weight relationship of Eq. 1. Based on the past studies, both f_c and m_c were maintained constant at 0.5% in the experiments. Based on the initial lime consumption test results, a constant lime content of 9%, C_l = 9%, was used in the lime stabilized specimens. Unconfined compression tests were carried out on a) untreated, b) fiber reinforced, c) mesh reinforced, d) lime stabilized, e) lime-fiber treated, and f) lime-mesh treated soil specimens. The fibers and meshes were assumed not to influence the values of MDD and OMC. All specimens were prepared at their respective MDD-OMC state. Thus the specimens in which no lime was used were prepared at dry unit weight, γ_d = 1,680 kN/m^3 and water content, w = 21% and all the lime stabilized specimens were prepared at γ_d = 1,560 kN/m^3 and w = 22%. The specimens without lime were tested immediately after preparation. The lime stabilized specimens were tested both immediately after preparation and after curing. The different curing periods, t, were 7, 14, 28, 56, 90, and 120 days.

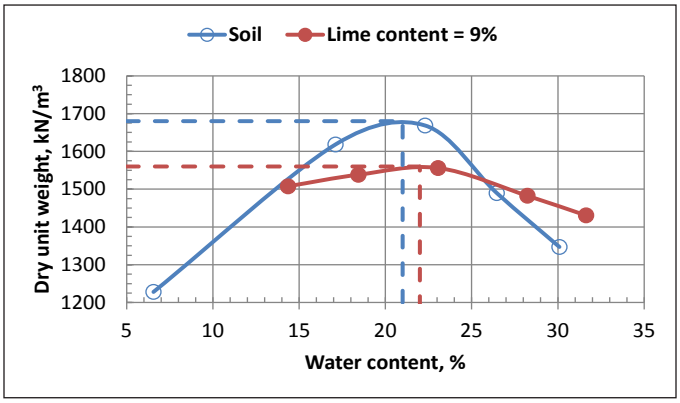


Figure. 3. Effect of lime on the compaction characteristics of soil

(a) (b)

Figure. 4. Scanning electron microscope images: (a) Fiber; and (b) Mesh

Preparation of Specimens. The soil was dried in the oven and then pulverized. The amount of soil, sufficient for the preparation of three identical specimens, was kept in a tray. For easy extraction of the specimens after compaction, lubricant was smeared on the inner surface of 35 mm diameter × 70 mm height cylindrical molds. Depending on how the soil was treated, the required quantities of lime, fiber, and mesh were measured and mixed with the dry soil evenly. Distilled water corresponding to the OMC of the treated specimens was measured. The dry mixture was mixed with the water evenly. The wet mixture was then placed immediately inside a polyethylene bag and sealed to prevent moisture loss. The specimen was prepared in three layers inside the mold; each layer was

compacted thirty times with a 10 kg spring compactor. Wet mixture for the first layer was placed inside the mold and compacted. The surface of the compacted layer was scarified before the material for the next layer was added. The compaction procedure was repeated for the second and third layers of the specimen. After the compaction of the last layer, the specimen was trimmed and extracted carefully from the mold. A sample extractor was used to prevent cracks and breakage of the specimen. Measurements for the weight and water content of the specimen were made. These readings were used to calculate the initial dry unit weight and water content of the specimens to determine the deviation from the target MDD and OMC values. The specimen was wrapped closely in a plastic wrap. A label was applied on the specimen for proper identification. Lime stabilized specimens were placed in a polyethylene bag and kept at laboratory temperature for curing.

Specimen Labeling Scheme. Since C_l was constant (9%) in all the lime stabilized specimens and f_c or m_c was constant (0.5%) in all the reinforced specimens, the labeling scheme mainly differentiated the type of reinforcement and the curing period. Figure 5 shows the specimen labeling scheme.

The first alphabet in the label indicates whether lime is used as binder or not (U = natural soil; L = lime stabilized). The second alphabet indicates the type of reinforcement (F = fiber; M = mesh). There is no second alphabet if no reinforcement is used. The next set of numbers indicates the curing period in days. Three specimens were prepared for each set of parameters to check the reproducibility of results. The last number indicates the number of specimen. For example, U0 – 2 stands for natural soil specimen number 2 with no curing. UM0 – 3 designates mesh reinforced natural soil specimen number 3 with no curing. LF14 – 1 indicates lime stabilized fiber reinforced specimen number 1 cured for 14 days.

Testing of Specimens. Unconfined compression tests were carried out on the specimens at the deformation rate of 1.27 mm/minute. On lime stabilized specimens, the tests were conducted immediately after their preparation and at the end of different curing periods. Specimens not treated with lime were tested immediately after their preparation. Before conducting the unconfined compression test, the weight of each specimen was recorded. At the end of the unconfined compression test, the water content of the specimen was determined. A total of 72 unconfined compression tests were carried out. SEM images were also obtained in selected specimens. Figure 6 shows the SEM image of a LM90 specimen. The fine grained soil was able to pass through the openings in the meshes and embed them completely in the soil.

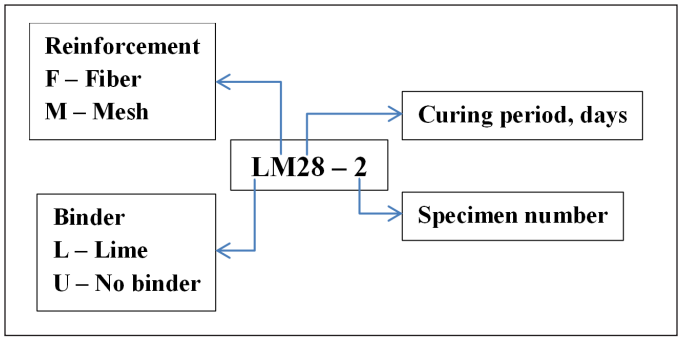


Figure. 5. Specimen labeling scheme

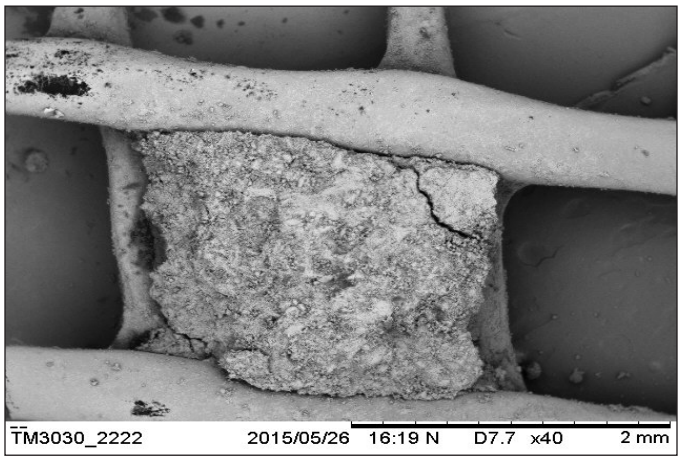


Figure. 6. Scanning electron image of LM90 specimen

RESULTS AND DISCUSSION

Initial State of the Specimens

Equation 6 was used to determine the deviations in the initial γ_d and w of the specimens from the target values of MDD and OMC. Figure 7 shows the results of the deviations.

$$\text{Deviation from target value} = \frac{\text{Actual value} - \text{Target value}}{\text{Target value}} \times 100 \qquad [6]$$

Figure 7 shows that the initial γ_d was mostly in the region of 98% of the target values. The average deviation was only −1.4%. However, the deviation in the initial w was relatively more; the average deviation was −5.7%. In one-third of the specimens the deviation in the initial w ranged from 6% to 14%. The reason for the higher deviation in

initial *w* was the difficulty in controlling the moisture loss during specimen preparation. The average deviation was small for the first specimen (3%) compared to the next two specimens. The average deviation was the largest for the third specimen (8.4%). Since, all the specimens had nearly identical initial γ_d, the differences in the initial *w* did not influence the experimental outcome significantly as explained later in the paper.

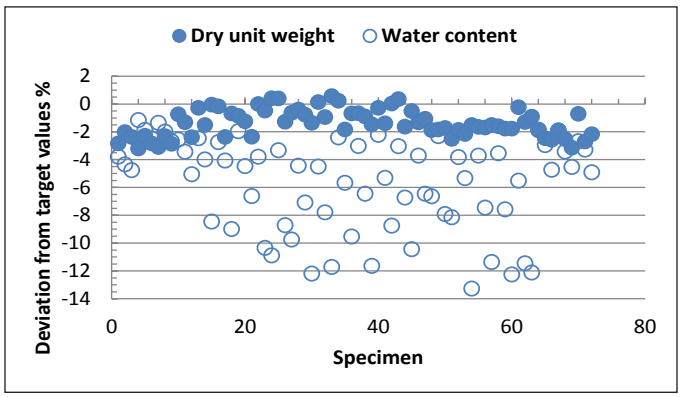


Figure. 7. Deviation in the initial γ_d and *w* from target values

Effect of Lime, Fiber, and Mesh on UCS

For a set of parameters the average of the UCS of the three particular specimens, UCS_{av}, was calculated. The variation of the UCS of each specimen from UCS_{av} was calculated using Eq. 7. Figure 8 shows the results of the variation in UCS for all the specimens from the corresponding UCS_{av} values.

$$\text{Variation from } UCS_{av} = \frac{UCS - UCS_{av}}{UCS_{av}} \times 100 \qquad [7]$$

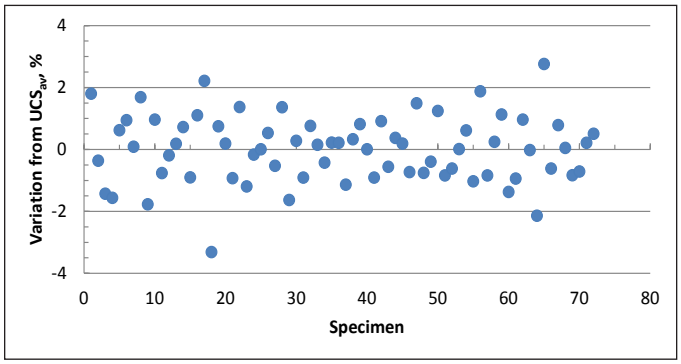


Figure. 8. Variation of UCS of the specimens from respective UCS_{av} values

The UCS of each specimen was very close to the corresponding UCS_{av} value; the variation was less than ±2%. Therefore, the UCS_{av} was representative of the UCS for any set of parameters. Further, the variation in the water content in the initial state of the specimens did not affect the UCS as the initial γ_d values of the specimens were nearly same.

Figure 9 shows the variation of UCS (= UCS_{av}) with time for the different methods of stabilization. The three horizontal lines at the bottom correspond to the soil mixtures in which no lime was used. These lines are shown as horizontal lines hypothetically assuming no increase in UCS with time in the natural soil. It is evident from the figure that both fibers and meshes increased the UCS of both the natural and lime stabilized soils; the increase was, however, more predominant in the former than in the latter. Further, the effect of lime on UCS was more prominent than of fibers and meshes. The strength of the lime stabilized mixtures increased up to about 28 days and thereafter remained almost constant. Figure 10 shows the percent increase in UCS of the natural soil due to different treatments. Both fibers and meshes increased the UCS of the natural soil by 30%. Santoni et al. (2001a) used very large size meshes, 51 *x* 102 mm, and reported that they contributed only to a small increase in strength. The present study indicates that small size meshes could increase the UCS of the soil better than large size meshes. Lime increased the UCS of the soil maximum by 70%. The combinations of lime and meshes, and lime and fibers increased the UCS of the soil maximum by 77% and 84%, respectively. Thus, the fibers performed better than the meshes.

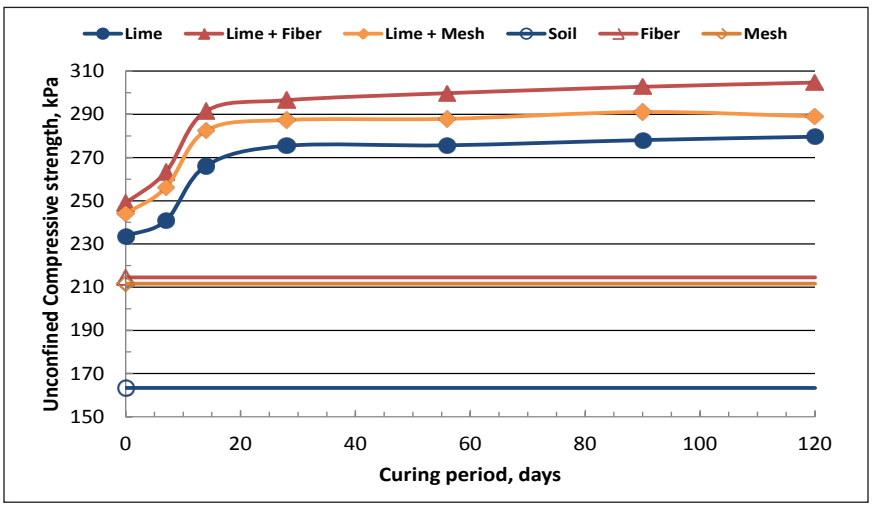


Figure. 9. Variation of UCS with curing time

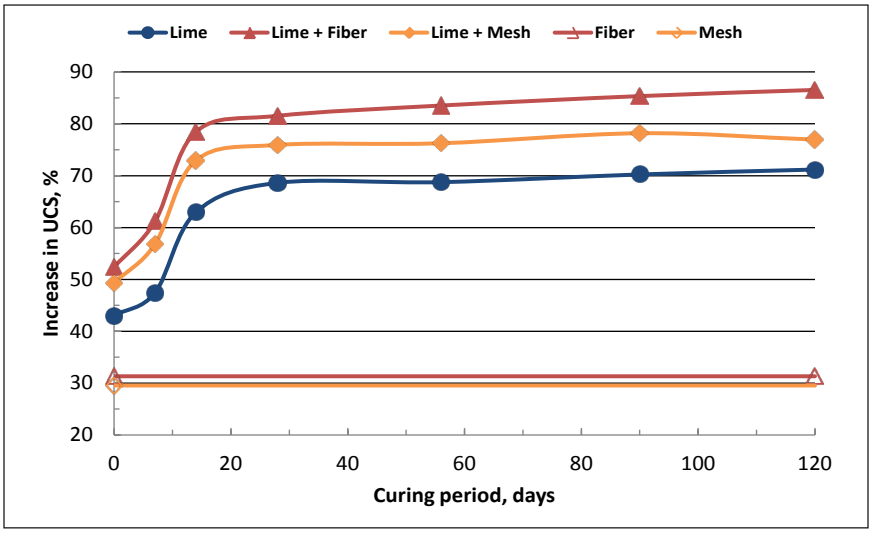


Figure. 10. Percent increase in UCS of the natural soil due to different treatments

Effect of Lime, Fiber, and Mesh on the Stress-Strain Behaviour

Figures 11 and 12 show how the fibers and meshes influenced the stress-strain behaviours of the natural soil and lime stabilized soil, respectively. In both cases, they increased the failure strain, ε_f, thus making the soil behaviour more ductile. For a set of parameters, the average ε_f of the three particular specimens, $\varepsilon_{f\text{-}av}$, was calculated. The variation of the ε_f of each specimen from $\varepsilon_{f\text{-}av}$ was calculated using Eq. 8. Figure 13 shows the variation in ε_f from the corresponding $\varepsilon_{f\text{-}av}$ values for all the specimens. As this variation was less than ±5% in about 90% of the specimens, $\varepsilon_{f\text{-}av}$ was the representative ε_f for any set of parameters.

$$\text{Variation from } \varepsilon_{f-av} = \frac{\varepsilon_f - \varepsilon_{f-av}}{\varepsilon_{f-av}} \times 100 \qquad [8]$$

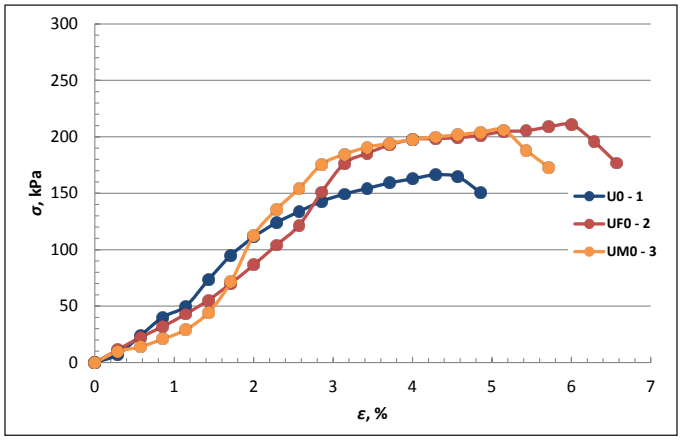


Figure. 11. Stress-strain curves of soil mixtures not stabilized by lime

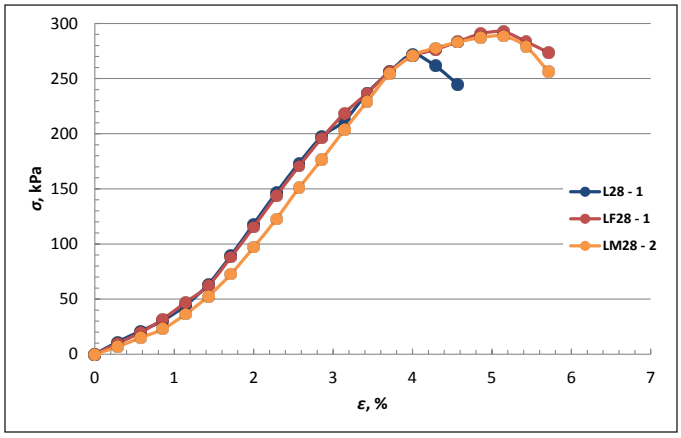


Figure. 12. Typical stress-strain curves of lime stabilized soil mixture specimens cured for 28 days

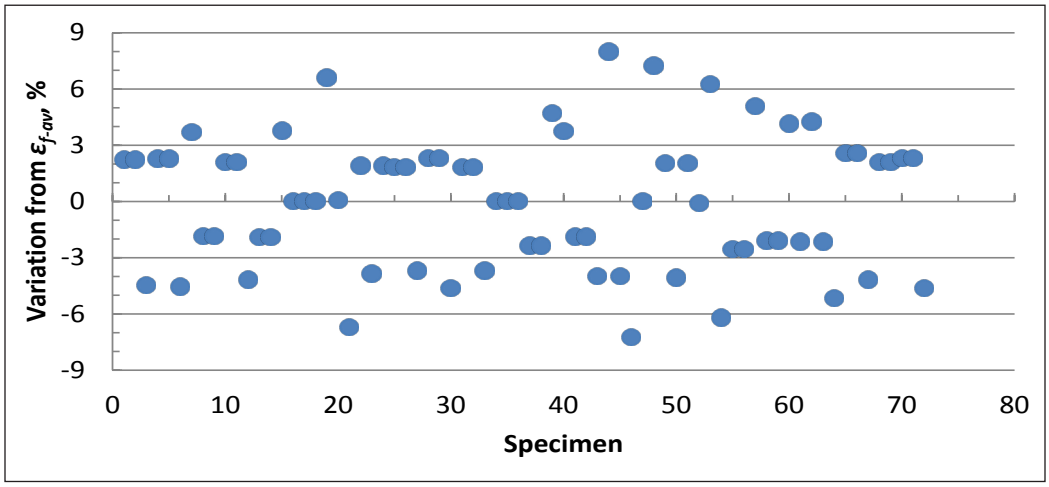


Figure. 13. Variation of ε_f of the specimens from respective $\varepsilon_{f\text{-}av}$ values

Figure 14 shows the variation of ε_f (= $\varepsilon_{f\text{-}av}$) with time for the different methods of stabilization. Generally, fibers performed better than meshes in increasing ε_f. In all lime stabilized mixtures, ε_f decreased as curing period increased and when soil was stabilized with lime only, ε_f became even less than that of the natural soil. Fibers and meshes increased the ε_f of lime stabilized soil, but eventually ε_f tended to be the same as that of the natural soil. Thus, fibers and meshes helped to regain the ductility lost due to addition of lime alone.

Effect of Lime, Fibers and Mesh Elements on Secant Modulus, E_{s50}

Equation 9 shows the expression for determining the secant modulus of a specimen corresponding to 50% of the peak axial stress, E_{s50}.

$$E_{s50} = \frac{0.5\sigma_{1f}}{\epsilon_{0.5\sigma_{1f}}} \quad [9]$$

In Eq. 9, σ_{1f} is the peak axial stress and $\epsilon_{0.5\sigma_{1f}}$ is the axial strain at one half the value of peak axial stress. The average values of E_{s50} for the soil-reinforcement mixtures U_0, U_{F0}, and U_{M0} not stabilized with lime were 7.4 MPa, 6.02 MPa, and 5.50 MPa, respectively. The values of E_{s50} of U_{F0} and U_{M0} were less than that of U_0. Thus, U_{F0} and U_{M0} were less stiff or more flexible than U_0. Flexibility induced by fibers and mesh elements in the stress-strain behaviour of the natural soil is evident. Figure 15 shows the variation of E_{s50} with time for the soil mixtures stabilized with lime and reinforcement. Comparison of Figs 9 and 15 shows that while the fibers and mesh elements increased the UCS of the lime stabilized soil, they decreased its secant modulus E_{s50}. Thus, fiber and mesh elements induced flexibility in the stress-strain behaviour of both natural and lime stabilized soils.

Mixing of Fibers and Mesh Elements

Fibers and meshes obtained from the HDPE insect net did not get entangled with each other. It was possible to separate and count the individual fibers and meshes in each specimen. On the average, there were 2,914,188 fibers or 564,161 meshes in one cubic meter of soil. Therefore, fibers and meshes made of materials similar to that in the present study could be used in practice to prepare soil-reinforcement mixtures in which fibers and meshes would be uniformly distributed.

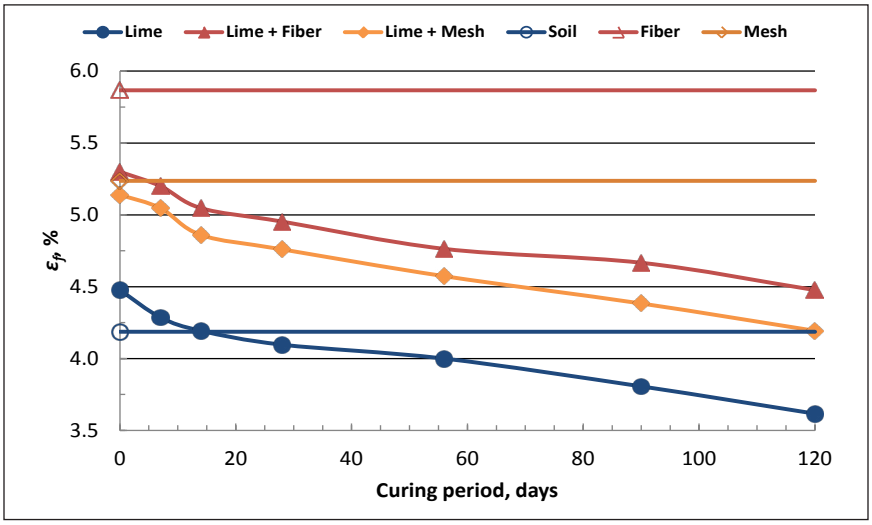


Figure. 14. Variation of ε_f with curing period

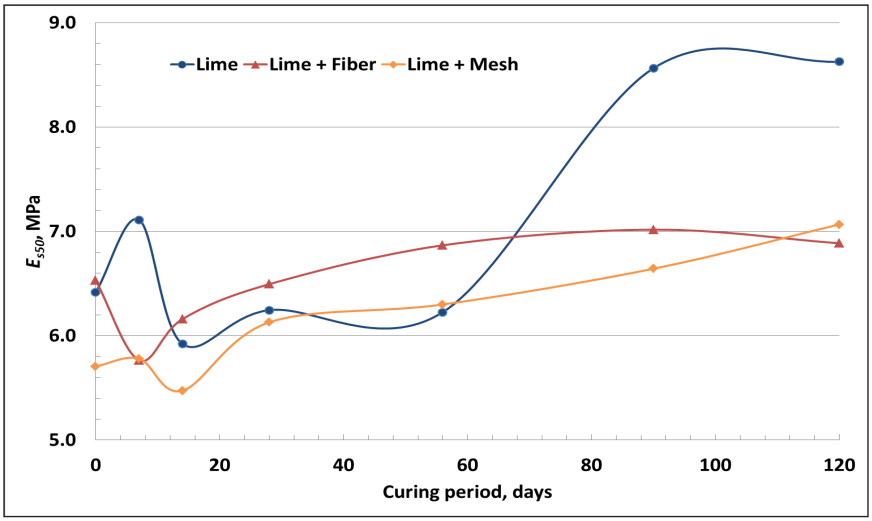


Figure. 15. Variation of secant modulus E_{s50} with curing period

CONCLUSION

Studies on fiber reinforcement began much earlier on coarse grained soils than on fine grained material. A binder was used together with fibers more commonly in the studies on fine grained soils than in granular soils. The general conclusions from a review of the past studies have been presented in the body of the paper for fiber reinforced granular soils and fine grained materials separately.

An experimental study was carried out on a fine grained soil using combinations of lime, fibers and meshes to improve the mechanical characteristics of the soil. Fibers and meshes were obtained from a HDPE insect net in view of economy and ease of mixing with the soil. The major conclusions from the study are as follows.

1. Fibers and meshes increased the UCS of both the natural and lime stabilized soils. Their influence was more significant on the natural soil than on the lime stabilized soil.
2. The effect of lime on UCS was more prominent than of fibers and meshes.
3. The UCS of the lime stabilized mixtures increased up to about 28 days and thereafter remained almost constant.
4. Fibers and meshes increased the failure strain, ε_f, and made the stress-strain behaviour of the soil more ductile. Generally, fibers increased the failure strain more than the meshes.
5. In all lime stabilized mixtures, ε_f decreased as curing period increased.

6. When soil was stabilized with lime only, ε_f became even less than that of the natural soil. Fibers and meshes increased the ε_f of the lime stabilized soil.
7. The secant moduli corresponding to one half of the peak axial stress, E_{s50}, are less for the fiber and mesh reinforced soil specimens than for the unreinforced specimens. This is due the ductility induced in the stress-strain behaviours of both the natural and lime stabilized soil specimens by the fibers and meshes.

ACKNOWLEDGEMENTS

The facilities provided by the Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan, Malaysia, for the experimental study are gratefully acknowledged.

REFERENCES

Ahmad, F., Bateni, F., & Azmi, M. (2010). Performance evaluation of silty sand reinforced with fibres. *Geotextiles and Geomembranes, 28*, 93-99.

Al-Refeai, T. (1991). Behaviour of granular soils reinforced with discrete randomly oriented inclusions. *Geotextiles and Geomembranes, 10*(4), 319-333.

Anggraini, V., Asadi, A., & Huat, B. K. (2015). Durability of lime treated soil reinforced by natural fiber under bending force. *International Journal of Geological and Environmental Engineering, 9*(12), 1400-1404.

Babu, G. L., Vasudevan, A. K., & Haldar, S. (2008). Numerical simulation of fiber-reinforced sand behavior. *Geotextiles and Geomembranes, 26*, 181–188.

Botero, E., Ossa, A., Sherwell, G., & Ovando-Shelley, E. (2015). Stress–strain behavior of a silty soil reinforced with polyethylene terephthalate (PET). *Geotextiles and Geomembranes, 43*, 363-369.

BS 1377. (1990). *Methods of test for soils*. British Standard Institution, London, U.K.

BS 1924. (1990). *Stabilised materials for civil engineering purposes*. British Standard Institution, London, U.K.

Butt, W. A., Mir, B. A., & Jha, J. N. (2016). Strength behavior of clayey soil reinforced with human hair as a natural fibre. *Geotechnical and Geological Engineering, 34*, 411–417.

Chauhan, M. S., Mittal, S., & Mohanty, B. (2008). Performance evaluation of silty sand subgrade reinforced with fly ash and fibre. *Geotextiles and Geomembranes, 26*, 429–435.

Chen, M., Long Shen, S. L., Arulrajah, A., Wu, H. N., Hou, D.W., & Xu, Y. S. (2015). Laboratory evaluation on the effectiveness of polypropylene fibers on the strength of fiber-reinforced and cement-stabilized Shanghai soft clay. *Geotextiles and Geomembranes, 43*, 515–523.

Consoli, N. C., Bassani, M. A. A., & Festugato, L. (2010). Effect of fiber-reinforcement on the strength of cemented soils. *Geotextiles and Geomembranes, 28*, 344–351.

Consoli, N. C., Corte, M. B., & Festugato, L. (2012). Key parameter for tensile and compressive strength of fiber-reinforced soil-lime mixtures. *Geosynthetics International, 19*(5), 409-414.

Consoli, N. C., Prietto, P. D. M., & Ulbrich, L. A. (1998). Influence of fiber and cement addition on behavior of sandy soil. *Journal of Geotechnical and Geoenvironmental Engineering, 124*(12), 1211-1214.

Consoli, N. C., Vendruscolo, M. A., Fonini, A., & Rosa, F. D. (2009). Fiber reinforcement effects on sand considering a wide cementation range. *Geotextiles and Geomembranes, 27*, 196–203.

Correia, A. A. S., Oliveira, P. J. V., & Custodio, D. G. (2015). Effect of polypropylene fibres on the compressive and tensile strength of a soft soil, artificially stabilised with binders. *Geotextiles and Geomembranes, 43*, 97-106.

Cristelo, N., Cunha, V. M. C. F., Dias. M., Gomes, A. T., Miranda, T., & Araújo, N. (2015). Influence of discrete fibre reinforcement on the uniaxial compression response and seismic wave velocity of a cement-stabilised sandy-clay. *Geotextiles and Geomembranes, 43*, 1-13.

Diambra, A., Ibraim, E., Muir Wood, D., & Russell, A. R. (2010). Fibre reinforced sands: Experiments and modelling. *Geotextiles and Geomembranes, 28*, 238–250.

Estabragh, A. R., Namdar, P., & Javadi, A. A. (2012). Behavior of cement stabilized clay reinforced with nylon fiber. *Geosynthetics International, 19*(1), 85-92.

Fatahi, B., Khabbaz, H., & Fatahi, B. (2012). Mechanical characteristics of soft clay treated with fibre and cement. *Geosynthetics International, 19*(3), 252-262.

Gray, D. H., & AI-Refeai, T. (1986). Behaviour of fabric versus fiber-reinforced sand. *Journal of Geotechnical Engineering Division, 112*(8), 804-820.

Hamidi, A., & Hooresfand, M. (2013). Effect of fiber reinforcement on triaxial shear behavior of cement treated sand. *Geotextiles and Geomembranes, 36*, 1-9.

Hoare, D. J. (1979). Laboratory study of granular soils reinforced with randomly oriented discrete fibers. *Proceedings of International Conference on Soil Reinforcement*, *1*, 47–52. Paris, France.

Hoover, J. M., Moeller, D. T., Pitt, J. M., Smith, S. G., & Wainanina. (1982). *Performance of randomly oriented, fiber-reinforced roadway soils – A laboratory and field investigation*. Iowa DOT Project HR-211, ERI Project 1427, ISU-ERl-Ames-83192, Report, College of Engineering Iowa State University, 1-311.

Kalumba, D., & Chebet, F. C. (2013). Utilisation of polyethylene (plastic) shopping bags waste for soil improvement in sandy soils. *Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering*, *1*, 3223-3226. Paris.

Kaniraj, S. R., & Gayathri, V. (2003). Geotechnical behavior of fly ash mixed with randomly oriented fiber inclusions. *Geotextiles and Geomembranes, 21*, 123–149.

Kaniraj, S. R., & Gayathri,V. (2006). Behavior of fiber-reinforced cement-stabilized fly ashes. *Journal of Testing and Evaluation, 34*(4), 290-297.

Kaniraj, S. R., & Havanagi, V. G. (2001). Behavior of cement-stabilized fiber-reinforced fly ash-soil mixtures. *Journal of Geotechnical and Geoenvironmental Engineering, 127*(7), 574-584.

Kim. Y. T., Kim, H. J., & Lee, G. H. (2008). Mechanical behavior of lightweight soil reinforced with waste fishing net. *Geotextiles and Geomembranes, 26*, 512–518.

Kumar, A., & Gupta, D. (2016). Behavior of cement-stabilized fiber-reinforced pond ash, rice husk ash-soil mixtures. *Geotextiles and Geomembranes, 44*, 466–474.

Legeay , G., Nguyen, Long, T., & Gugen, Y. (1972). *Etude de la terre armee a l'appareil triaxial* [A study of lever arm with triaxial apparatus]. Rapport de Recherche RRO 17. France, Paris: Laboratoire Central des Ponts et Chaussees.

Li, C. (2005). *Mechanical response of fiber-reinforced soil* (Doctoral thesis). The University of Texas at Austin, U.S.A.

Li, C., & Zornberg, J. G. (2013). Mobilization of reinforcement forces in fiber-reinforced soil. *Journal of Geotechnical and Geoenvironmental Engineering, 139*(1), 107-115.

Maheshwari, K., & Solanki, C. H. (2009). Behaviour of fiber reinforced soil. *Australian Geomechanics, 44*(4), 65-74.

McGown, A. (1978). Effect of inclusion properties on the behavior of sand. *Geotechnique, 28*(3), 327-346.

Mirzababaei, M., Miraftab, M., Mohamed, M., & McMahon, P. (2013). Unconfined compression strength of reinforced clays with carpet waste fibers. *Journal of Geotechnical and Geoenvironmental Engineering, 139*(3), 483-493.

Olgun, M. (2013). Effects of polypropylene fiber inclusion on the strength and volume change characteristics of cement fly ash stabilized clay soil. *Geosynthetics International, 20*(4), 263-275.

Oliveira, P. J. V., Correia, A. A. S., Teles, J. M. N. P. C., & Custódio, D. G. (2015). Effect of fibre type on the compressive and tensile strength of a soft soil chemically stabilised. *Geosynthetics International, 23*(3), 171-182. doi: http://dx.doi.org/10.1680/jgein.15.00040

Park, S. S. (2009). Effect of fiber reinforcement and distribution on unconfined compressive strength of fiber-reinforced cemented sand. *Geotextiles and Geomembranes, 27*, 162–166.

Park, T., & Tan, S.A. (2005). Enhanced performance of reinforced soil walls by the inclusion of short fiber. *Geotextiles and Geomembranes, 23*, 348–361.

Pino, L. F. M., & Baudet, B. A. (2015). The effect of the particle size distribution on the mechanics of fibre-reinforced sands under one-dimensional compression. *Geotextiles and Geomembranes, 43*, 250-258.

Plé, O., & Lê, T. N. H. (2012). Effect of polypropylene fiber-reinforcement on the mechanical behavior of silty clay. *Geotextiles and Geomembranes, 32*, 111-116.

Qu, J., & Sun, Z. (2016). Strength behavior of Shanghai clayey soil reinforced with wheat straw fibers. *Geotechnical and Geological Engineering, 34*(2), 515-527. doi: 10.1007/s10706-015-9963-8.

Rao, S. V. K., & Nasr, A. M. A. (2012). Laboratory study on the relative performance of silty-sand soils reinforced with linen fiber. *Geotechnical and Geological Engineering, 30*, 63-74.

Sadek, S., Najjar, S. S., & Freiha, F. (2010). Shear strength of fiber-reinforced sands. *Journal of Geotechnical and Geoenvironmental Engineering, 136*(3), 490-499.

Santoni, R. L., & Webster, S. L. (2001). Airfields and road construction using fiber stabilization of sands. *Journal of Transportation Engineering, 127*(2), 96–104.

Santoni, R. L., Tingle, J. S., & Webster, S. L. (2001). Engineering properties of sand-fiber mixtures for road construction. *Journal of Geotechnical and Geoenvironmental Engineering, 127*(3), 258–268.

Shao, W., Cetin, B., Li, Y., Li, J., & Li, L. (2014). Experimental investigation of mechanical properties of sands reinforced with discrete randomly distributed fiber. *Geotechnical and Geological Engineering, 32*, 901-910.

Tang, C., Shi, B., Gao, W., Chen, F., & Cai, Y. (2007). Strength and mechanical behavior of short polypropylene fiber reinforced and cement stabilized clayey soil. *Geotextiles and Geomembranes, 25*, 194–202.

Texas Department of Transportation. (2005). *Guidelines for modification and stabilization of soils and base for use in pavement structures*. pp. 1-25.

Verma, B. P., & Char, A. N. (1978). *Triaxial tests on reinforced sand*. Proceedings of the Symposium on Soil Reinforcing and Stabilising Techniques (pp. 29-39). Sydney, Australia

Yang, Z. (1972). *Strength and deformation characteristics of reinforced sand* (PhD Dissertation). University of California, U.S.A.

Yetimoglu, T., & Salbas, O. (2003). A study on shear strength of sands reinforced with randomly distributed discrete fibers. *Geotextiles and Geomembranes, 21*, 103–110.

Yi, X. W., Ma, G. W., & Fourie, A. (2015). Compressive behaviour of fibre-reinforced cemented paste backfill. *Geotextiles and Geomembranes, 43*, 207-215.

Pertanika J. Sci. & Technol. 26 (4): 1571 - 1590 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Automatic Annotation of Educational Videos for Enhancing Information Retrieval

Poornima, N.* and Saleena, B.

School of Computing Science and Engineering, VIT Chennai, Tamil Nadu - 600 127, India

ABSTRACT

Educational videos are one of the best means of imparting knowledge to the users/learners. Videos can convey information in an effective and interesting manner. These videos can be accessed through online or from stored repositories using queries. Search queries play important role in the retrieval. Whenever a user gives an ambiguous query, the search engine may produce irrelevant results. Thus a lot of time is being spent by the users in retrieving the relevant videos. In order to improve the probability of retrieving relevant results, semantic web technologies are applied. This paper aims to extract keywords from the videos and to find the association between the extracted terms. The associated terms are arranged based on their frequency of occurrences. These terms are used to annotate the video automatically, which in turn improves the retrieval of more relevant videos. An ontology is created by experts based on the e-learning video domain. Videos are grouped based on the keywords and on domain ontology, which also helps in enhancing the retrieval results. Videos containing text are only considered for processing.

Keywords: Annotation, e-learning, ontology, semantics, Term Frequency Inverse Document Frequency (TF-IDF), video retrieval

INTRODUCTION

Use of technology and internet in the modern educational system makes learning more interactive and interesting. Nowadays universities and educational institutions follow various techniques to improve the student's learning skills, for example: flipped classroom, collaborative learning, differential learning and virtual classrooms. In this modern educational system, study

ARTICLE INFO

Article history:
Received: 07 August 2017
Accepted: 27 June 2018
Published: October 2018

E-mail addresses:
poornima.n2014@vit.ac.in (Poornima, N.)
saleena.b@vit.ac.in (Saleena, B.)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

materials are delivered to the users in various modes such as document, presentation slides, audios or videos. These materials can be delivered directly to the users or it can be stored in a repository, so that users can access materials through querying.

After the invention of video sharing websites, educational videos have become popular among learners. A survey showed that there was tremendous increase in the amount of educational videos uploaded in YouTube (Che & Lin, 2015). YouTube has introduced a separate channel YouTube EDU for learners and educators. Large number of educational videos makes the search process tedious and time consuming. In case of text document search, entire document is analysed and results are displayed based on the user query. Unlike text document search, videos are retrieved only based on the annotation given to the video without any analysis on the video contents. In general, videos are annotated manually by the authors of video. Manual annotation consumes more time and there is also a possibility of giving improper keyword annotation, just to attract the users.

The central goal of this research is to automate the annotation process by analysing the video contents and to provide semantic meaning to the keywords to enhance the information retrieval. Text from the videos are captured using Tesseract Optical Character Recognition (OCR) which helps in analysis of the video contents. Using WordNet, semantic meanings and relationships of extracted words are found. Term frequencies for all the extracted words from video and WordNet are identified and arranged based on frequent occurrence. In addition, domain ontologies are created for all the categories of videos available in database. According to domain ontologies, highly occurring words are grouped along with its videos. If the user query is related to any of the video group, then the entire group will be retrieved.

Major contributions of the paper include:

- Automatic video content Extraction: Visual analysis techniques such as keyframe extraction and text extraction from video helps in analysing the video contents. Textual features are identified and extracted using Video OCR technique.
- Filling semantic gap between query and retrieved videos: Keywords are identified from the videos. The related terms of the keywords are retrieved using WordNet. The keywords along with the related terms and the videos were clustered based on the domain ontology, which was helpful in improving the retrieval results.

Section 2 discusses some of the open source coursewares and its major issues during retrieval. Section 3 discusses the related work carried out in this field. Section 4 describes the methodology for information retrieval using semantic web technologies. Section 5 discusses experimentation setup, results and performance evaluation metrics. Finally section 6 concludes with future directions for this research.

Learning Issues in E-Learning Coursewares

Many open source coursewares by major universities allow students to watch and download course materials anytime and anywhere from the world. Massachusetts Institute of Technology: MIT (MITOPENCOURSEWARE), Harvard University (Harvard Open Learning Initiative) and University of California (UC Irvine, OpenCourseWare) are some of the open coursewares provided by popular universities.

Population of engineering students in India outnumbers ("NPTEL Frequently Asked Questions", 2018) every other country. NPTEL is a curriculum building exercise aims to create open source contents for major science and engineering courses. NPTEL projects are funded by Government of India and it is used by most of the Indian engineering students. NPTEL website does not allow keyword based search. Search on NPTEL website can be done only based on the course name and professor name. NPTEL courses can be filtered by discipline, content type (like subject) and institutions (Figure 1). Searching a topic "Properties of transaction" from NPTEL responds with no result (Figure 1). Database Design course video contains content on "Transaction Properties" (Figure 2). Even Database Design course syllabus contains topic "Properties of transactions" (Figure 2).

In case of MITOPENCOURSEWARE, users can browse through a topic by selecting a topic, subtopic and specialty (Figure 3). After choosing from the list of topic, sub topic and specialty, results are given. Results for the topic "Data mining" from MITOPENCOURSEWARE are shown in figure (Figure 3).

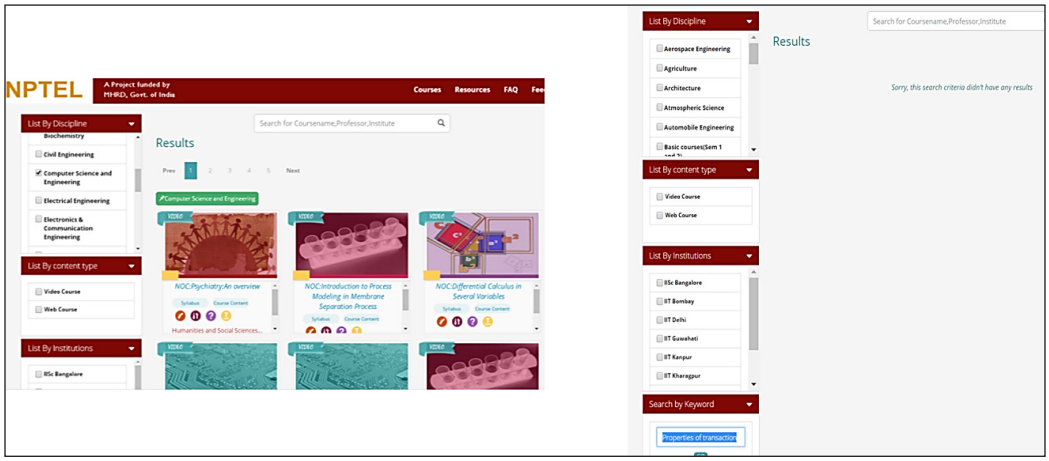


Figure 1. Screenshot showing different search facilities available in NPTEL (Selection of Discipline -> Civil Engineering, Computer Science and Engineering, Electronics & Communication Engineering and so on, content type -> Video Course and Web Course and Institutions -> IISC Bangalore, IIT Kanpur, IIT Madras and so on) and Result of searching topic "Properties of Transaction" through Keyword Search option in NPTEL

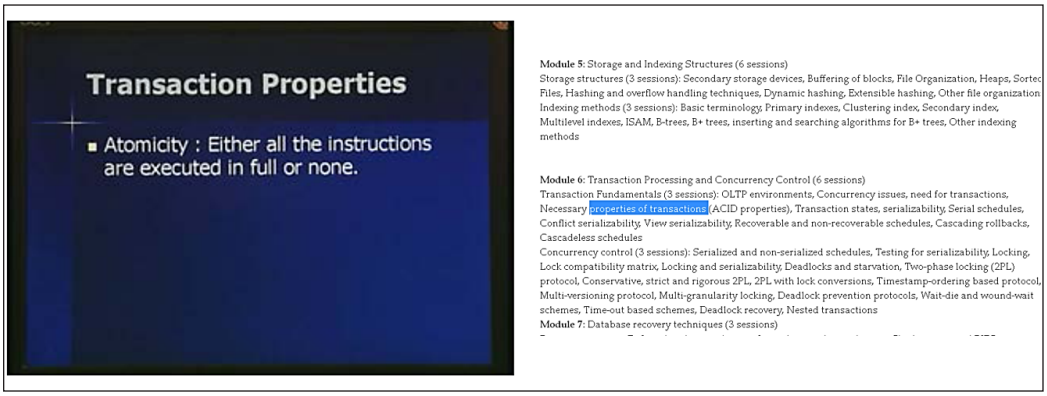


Figure 2. Syllabus and Video frame of NPTEL course 'Database Design' showing 'Transaction Properties' in it

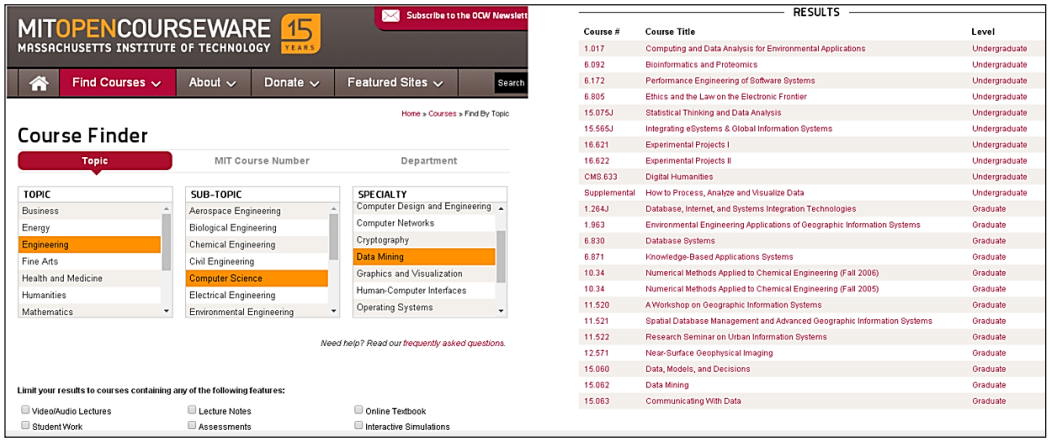


Figure 3. Course Finder of MITOPENCOURSEWARE (showing selection of Topic, Sub-Topic and Speciality) and Search result of course 'Data Mining' in MITCOURSEWARE provides irrelevant course results

Similar challenges are also encountered in Harvard Open Learning Initiative and UC Irvine open coursewares. Search by instructor, keyword and course are the possible ways for searching in Harvard Open Learning Initiative. Keyword search results are not more relevant to the search query (Figure 4). Even though University of California (UC Irvine, OpenCourseWare) allows you to search the content through keywords, search results are completely irrelevant to the user query (Figure 5). Search through type and categories are also possible in UC Irvine OpenCourseWare.

In the above open coursewares, retrievals are inefficient because videos are not analysed based on the contents. Learner can fetch the materials only based on limited options such as selection through department/school, category/subject, topic/subtopic and so on. Videos are mostly annotated with random keywords which are totally inappropriate and leads

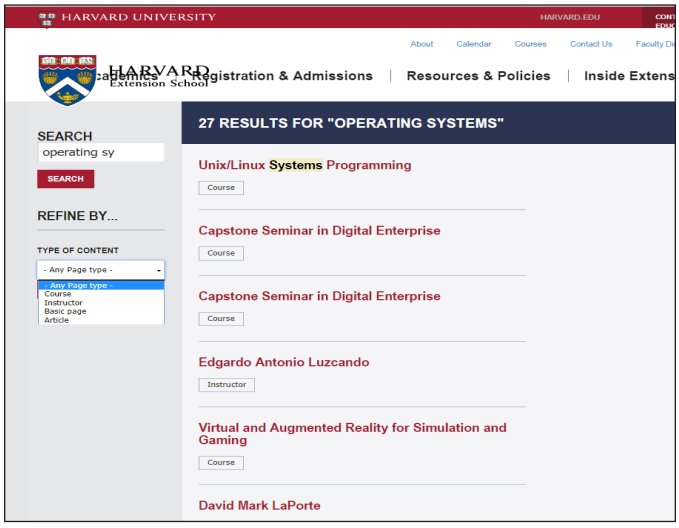


Figure 4. Search result of course 'Operating Systems' in Harvard Open Learning Initiative showing irrelevant course results

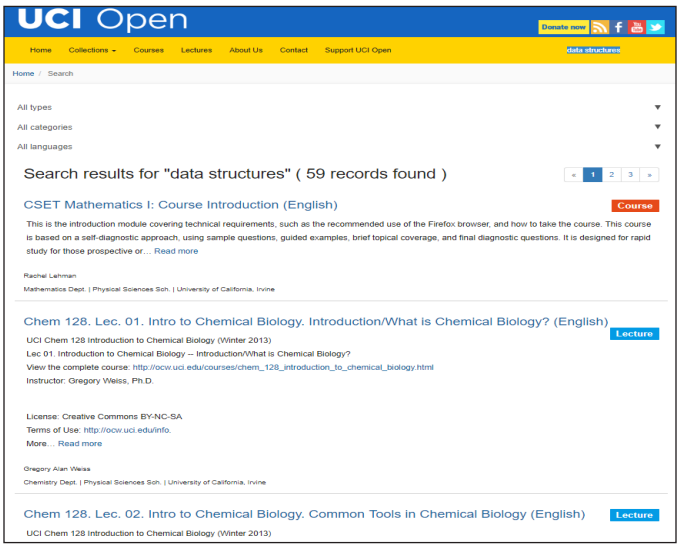


Figure 5. Search result of course 'Data Structures' in University of California (UC Irvine, OpenCourseWare) showing irrelevant course results

to irrelevant results. Some of the reputed journal papers (Balasubramanian et al., 2016; Muralikumar et al., 2016) have mentioned several open standard coursewares. They have clearly mentioned that the coursewares largely dependents on the tags, annotations and limited user-provided data for video retrieval. Lectures on these coursewares are presented with topic based segments, however, the structure and the organization of lectures is a result of manual processing. The search supported by most of these systems is mainly

occurrence based or tag-based, where the occurring search terms are highlighted in the transcripts. Some systems allow for navigating directly to the place where the search terms occur. Since annotations are given by users, there is a possibility for random keywords.

Annotating video with appropriate keywords from the video visuals will improve the accuracy which in turn improves precision. The search results can be improved by annotating the videos with the relevant keywords, which in turn will improve the search results.

Related Work

This section discusses about the various video retrieval systems using the video features and different methods for semantic video annotation.

Videos are segmented into shots. From each shot, shot level objects are selected by the user (El-Khoury et al., 2013). To identify objects in the shots, concept detectors are trained using classification algorithms such as k-nearest neighbour (k-NN), Support Vector Machine (SVM) or decision tree. Object features are extracted using Scale invariant feature transform (SIFT) descriptors. User chosen object are tracked by feature extraction methods. Annotation for each shot is done by concept detectors. Ten key frames are selected in algorithm using k-means clustering (Ravinder & Venugopal, 2016). Texture, edge and motion features are combined from all the ten key frames to represent feature vector. Feature vectors of query video and videos from video repository are compared for finding relationship between them using Euclidean distance measure. Resultant videos with less Euclidean distance are retrieved. However there is no defined methodology for filling the semantic gap between the content and the retrieved results.

Character regions are identified using Maximally Stable Extremal Regions (MSERs) based on the stroke width, letter height and width, character spaces (Wattanarachothai & Patanukhom, 2015). Text candidates are then classified according to their height and width. Tesseract OCR is applied to recognize the characters. Super-Fast Event Recognition system combines features of static visual descriptor, motion descriptor and audio descriptor (Jiang et al., 2015). All the feature descriptors are converted into bag of words representation. Then SVM kernel classifier is used to classify the events based on each feature. Middle level representation is created in Li's system to bridge the semantic gap between the low level features and high level features from the videos (Li et al., 2015). Middle level representations are built using Latent Dirichlet Allocation (LDA). To improve the system further and to reduce the computational cost, SIFT descriptors from LDA and fisher vectors are combined.

The work presented by Viana and Pinto (2017) proposed a video content annotation that used the concepts of crowdsourcing and gamification to collect metadata. To enhance the search and access, metadata were linked to specific time stamps. Semantic concepts

were identified through crowdsource tag-based dictionaries instead of standard dictionaries and thesaurus. Such semantic concepts lacks validation since system depends upon the external resources.

CourseMapper (Chatti et al., 2016) is an annotation platform that enables learners to collaborate and interact with video lecture using visual learning. Annotation editor allows user to add annotation to the viewed videos. These annotations are used for Visual learning. Visual learning methods are based on Annotation Maps and Heatmaps. AnnotationMap overlays stacks of annotation windows within the given timeline. To minimize the user distractions and to simplify the visual seeking for annotations, the cue points are marked in yellow color. This notifies the user that this portion of the video timeline has a larger number of annotations and most likely contains interesting information. Heatmap highlights the most viewed parts of video with warm colours such as orange and red, and less viewed parts are usually highlighted with cold purple and blue colours. Using this user can easily find the most interesting part of the video. Heatmap also records and displays the view count. This approach depends on the manual annotation of the authors.

In Kravvaris system, speech transcript for each video is collected from the repository (Kravvaris et al., 2015). These transcripts and the search query are converted into vectors. Cosine similarity between these two vectors is found. In addition to that, social weight to each video is added. That is likes and dislikes given by the registered YouTube users are taken into account. Based on the cosine similarity and the (like + dislike) values, videos are ranked.

Semantic annotation platform used by Xu & Mei (2015) enabled user to semantically annotate videos using vocabularies defined by traffic events ontologies. At first video annotation ontology is designed by following the traffic law which is machine understandable data. Description for video resources are given by annotator using those traffic field vocabularies. Semantic relationship between the annotations are used for management of annotated videos. However here annotation is carried out manually by the authors. Correlated Naive Bayes CNB classifier combines the methods of correlation and naïve bayes to retrieve relevant videos using the visual contents (Poornima & Saleena, 2018).

In this article, we define an approach that extracts visual contents of lecture video for automatic and semantic annotation. In addition, proposed approach does not rely on external resources such as social media, crowdsourcing, etc. for semantic video retrieval.

Educational Semantic Content Video Retrieval

Educational videos contain images, colourful illustrations, audio and many more. Most of the educational videos contain text as major feature (Balasubramanian et al., 2016; Li et al., 2015; Yang & Meinel, 2014). Analysing those text data and using it for annotation will

improve the search results. This automatic annotation will reduce the problem of author annotation and the time consumed for it. From the user's perspective, the user may not have clear idea on how to search a topic with proper query. So the user queries will be ambiguous and it will not match with the subjective contents of the video. This problem will be solved by giving meaning and relationships in the contents of video using WordNet ontology (http://wordnet-rdf.princeton.edu/ontology). WordNet ontology act as a thesaurus which groups English words based on synonyms. WordNet represents number of relationships between the members of WordNet. Workflow of the proposed work is shown in Figure 6. To annotate a video automatically requires complete analysis of the video. Analysis of the video needs image processing operations to extract the content and then data mining operations are needed for further processing of the content.

The process of semantic and automatic annotation based on video content involves the following two steps:

(i) Extraction of video contents

(ii) Generation and grouping of semantic keywords.

First step (Extraction of video contents) of the proposed work carries out image processing operations (Algorithm steps 1-3) such as selection of keyframe from video and text extraction from keyframe. Second step (Generation and grouping of semantic keywords) carries out datamining operations (Algorithm steps 4-7) such as detection of most frequent words, identification of semantic relationships and grouping of semantically related words & videos.

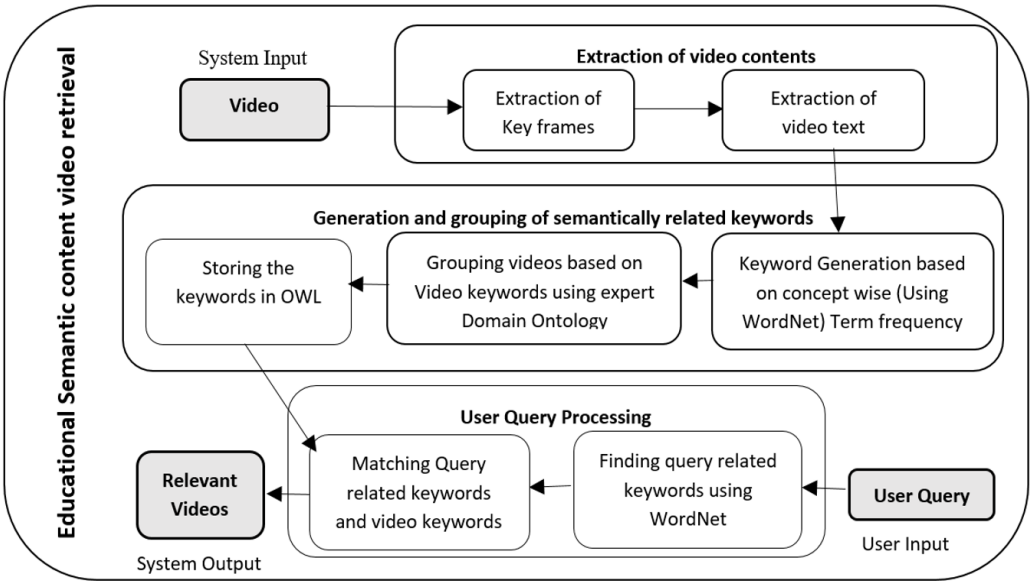


Figure 6. Educational Semantic video retrieval – Block Diagram

Extraction of Video Contents

Analysing a video content involves breaking the video into frames which is the basic element of a video. Unlike other videos, educational video contains text as major source of information. For the experimental setup, as of now only JPEG video formats are taken, but it can be extended for all other video formats. Extracting text contents from an educational video requires (i) Extraction of key frames and (ii) Extraction of text from videos.

Extraction of Key Frames

Video is made up of collection of frames. Frames are arranged in a temporal order to get a sequential flow. Key frames are the collection of frames which represents all the major elements of video. Key frames gives summary/abstraction of a video. Checking the transition change is the major task in finding the key frame. Several methods have been discussed in the literature for choosing the key frames, they are

(i) Reference frame: Reference frame is generated manually and then each video frame is compared with the reference frame to find the key frame (Ferman & Tekalp, 2003). Accuracy of the keyframe selection solely depends on the accuracy of reference frame selection.

(ii) Sequential comparison: Current frame and previous frame are compared based on pixel value. If there is much dissimilarity then the current frame will be taken as next key frame (Zhang et al., 2003). This method is very simple but there may be a repetition of same key frame since key frame analysis is carried out only based on local properties.

(iii) Clustering: Frames are clustered into groups and the frame which is nearer to the cluster center is taken as key frame. Accuracy of key frame selection depends on accuracy of clustering method used and its results (Yu et al., 2004). And also setting the number of key frames/cluster for grouping is difficult. Advantage over other methods is that it reflects global features of video.

Reference frame and sequential frame comparison methods uses pixel difference as key factor. Hence keyframe extraction techniques can be grouped into two approaches, pixel differencing approach or clustering approach.

Pixel comparison consists of finding the distance between the pixel values of consecutive frames. Pixel based methods compare a specific pixel in one frame with a corresponding pixel in a successive frame. Frames are converted into grayscale/black and white images before the comparison. Euclidean distance is defined in equation 1.

$$d\ (frame1, frame2) = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}\ \left(frame1(i,j) + frame2(i,j)\right)^2} \quad [1]$$

where d is the distance measure. Salt and pepper noise in the video may affect the keyframe selection accuracy while using pixel level comparison (Yang & Meinel, 2014) method.

Histogram based methods are alternative to pixel-based methods (Janwe & Bhoyar, 2016). Histogram gives the color distribution of the image. Successive similar frames will contain approximately the same color information and will have a similar histogram. Histogram difference between two frames is calculated using equation 2.

$$HFD = \{Histogram\ of\ 1^{st}\ frame - Histogram\ of\ 2^{nd}\ frame\} * Number\ of\ Gray\ levels \quad [2]$$

If HFD of particular frame is more than the threshold value, then that frame will be taken as key frame. Threshold can be calculated using equation 3. Drawback is that images with similar histograms may have different visual appearance.

$$Threshold = Mean\ Deviation + (a * Standard\ Deviation) \quad [3]$$

where a is a constant.

Lecture video contents such as text lines, images, and tables, can be taken as connected components of an image (Yang & Meinel, 2014). So connected components method is used in this proposed work to identify the key frames (Figures 7 - 8). Two pixels are said to be connected, if there is a path from one pixel to the other i.e., both pixels share same intensity value. Connected components can be identified by analysing the image from left to right and top to bottom. Connected component can be easily determined by giving a pixel particular label value. For example: Labelling of pixel p can be done through following information:

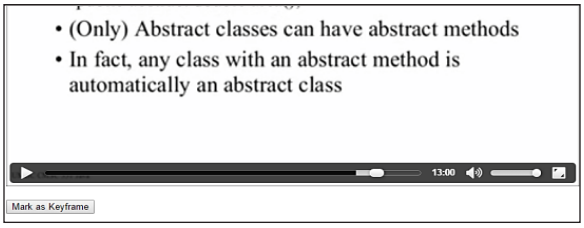


Figure 7. Selection of Key Frame from any online video

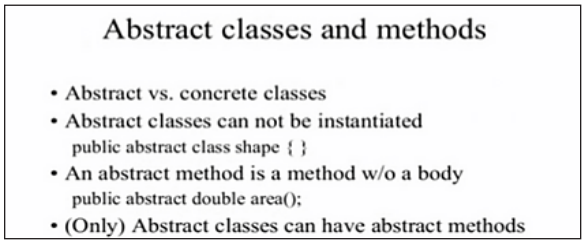


Figure 8. Selected Key frame taken as image for further text extraction

Step 1: If all four neighbors of a pixel p are 0, assign a new label to pixel p, else
Step2: If only one neighbor of a pixel p has V={1}, assign its label to p, else
Step3: If more than one of the neighbors have V = {1}, assign one of the labels to p and make a note of the equivalences.

Extraction of Text from Videos

Lecture contents closely depends on the text in lecture slide. These texts help in retrieval task for automation. Text extraction from video includes text detection and text recognition. Text detection refers to the presence of text in the frames. Text recognition is the process of converting text present in images into machine readable data.

Naturally, text has some properties such size, similar pattern, more interest point, high contrast than the background, connectedness and many. Text detection algorithms focus on these properties to detect text. Presence of text in frames is identified using connected components methods. Text lines are the major content of the video frame which are used for finding the connectedness. Tesseract is an open source OCR used to extract the video text and are converted into edited format for using it in next step. Key frames are converted into black and white images as a pre-processing for Tesseract. The procedure for conversion is as follows: step 1. Convert the colour images into binary/black and white image, step 2. Identify blobs and character lines, step 3. Match the character lines with pre-trained character set to find the character. For implementation, Java Tesseract (Figure 9) is used to recognize the text characters from the image frame. Then the extracted text are stored in an editable format (i.e., doc or txt) for further processing. Accuracy of Tesseract OCR are enhanced by using dictionaries.

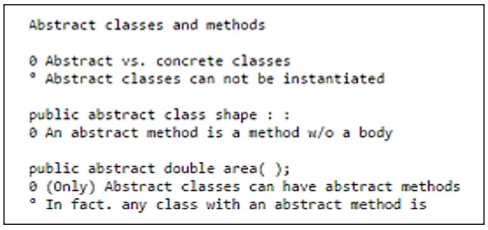


Figure 9. Extracted text in editable format from the selected video key frame

To extract information from a bulk data, which will lead to some meaningful pattern or knowledge, pre-processing plays very important role. Pre-processing includes case folding, stopwords removal, tokenization, Parts-Of-Speech (PoS) tagging, stemming and lemmatization. Varieties of capitalization may affect the processing, most common approach is to reduce all the words into lower cases (case folding). Most of the words in the sentence are the connecting parts rather than showing the subjects, objects or intent. Those words can be removed by comparing it with list of stopwords. Some examples of stopwords are

'of', 'the', 'an'. Tokenization describes splitting paragraphs into sentences, or sentences into individual words. Sentences can be split into individual words and punctuation through splitting across white spaces.

A Part-Of-Speech Tagger (POS Tagger) reads text and assigns parts of speech to each word, such as noun, verb, and adjective. The goal of both stemming and lemmatization is to reduce inflectional forms and sometimes derivationally related forms of a word to a common base form. However, the two words differ in their flavour. Stemming refers to a crude heuristic process that chops off the ends of words in the hope of achieving this goal correctly most of the time, and often includes the removal of derivational affixes. Lemmatization refers to doing things properly with the use of a vocabulary and morphological analysis of words, normally aiming to remove inflectional endings only and to return the base or dictionary form of a word, which is known as the lemma. For example: The stemmed form of analysis is: analysis and the lemmatized form of leaves is: analysis.

Generation and Grouping of Semantically Related Keywords and Videos

Relationship between the text extracted from the videos is found using WordNet. Finally frequency of terms is calculated for arranging the videos.

Relationship between Extracted Texts Using Wordnet

Each word from the extracted text is given meaning and relationships using WordNet. WordNet is a large lexical database of English words that are connected together by their semantic relationship. WordNet acts as both dictionary and thesaurus. Using WordNet ontology (Figure 10), meaning and relationships between the words in the video contents are found. In general, WordNet ontology has some flaws. If the dictionary is domain based, then it may be exactly suiting to our needs.

Algorithm:

Input: User Query and Video Database

Output: Set of relevant lecture videos

Step 1: Let us consider a video database D, which is a collection of lecture videos.

$$D = \{ V_1, V_2, \ldots, V_D, \ldots, V_n \} \quad [4]$$

where $V_1, V_2, \ldots V_n$ are the individual videos in the collection and n is the number of videos.

Step 2: Every individual video contains n number of frames in it. Let us consider V_D that has K number of keyframes.

$$V_D = \{V_D^1, V_D^2, \ldots, V_D^F, \ldots, V_D^K\} \quad [5]$$

where $V_{D,}^{1}$ V_{D}^{2}, .. are number of frames in D^{th} video and K represents total number of keyframes.

Step 3: Let us consider a video frame V_{D}^{F} that contains the textual features. The number of keywords generated from the frame V_{D}^{F} is,

$$W = \{ w_1, w_2, \ldots, w_m \} \quad [6]$$

where V_{D}^{F} is the F^{th} frame in the D^{th} video, w_1, w_2, . . are the keywords generated from the key frame $V_D{}^F$ and m is the total number of keywords extracted from the key frame V_{D}^{F}. Extraction of the keywords from the key frames follows the OCR technology, which uses the tesseract classifier.

Step 4: Each keyword carries some set of semantic words. Semantic words are the words which has similar synonym.

$$WS_1 = \{ ws_1^1, ws_1^2, \ldots ws_1^t \} \quad [7]$$

where $ws_1{}^1$, $ws_1{}^2$, . . , $ws_1{}^n$ are the semantic words extracted from the word w_1 and $ws_1{}^t$ represent the t^{th} semantic word generated from w_1 word. These semantic words are extracted from the keywords using WordNet Ontology.

Step 5: Term Frequency-Inverse Document Frequency (TF-IDF) is used to measure the importance of a keyword. tf(t,d) is the frequency of term t in document d.

$$idf(t) = log_2 \left(\frac{D}{df(t)} \right) \quad [8]$$

where df(t) is the document frequency and D is total number of documents in the domain corpus. TF−IDF for term t in document d is

$$tf - idf_{(t,st),d} = tf_{(t,st),d} * idf_t \quad [9]$$

where t is term and st is semantic term. TF-IDF helps in differentiating domain-specific terms and highly generic terms.

Step 6: Based on the words, semantic words and frequency, words are clustered into group. Grouping depends on the minimum distance between the clusters.

Step 7: When query arrives, the classifier matches query with clusters. Cluster with the maximum probability gets retrieved.

```
> findAssocs(dtm, "data", corlimit=0.6)
          data
analysi   1.00
base      1.00
becom     1.00
call      1.00
comput    1.00
consist   1.00
follow    1.00
ieee      1.00
issu      1.00
messag    1.00
micro     1.00
part      1.00
predict   1.00
process   1.00
tool      1.00
use       1.00
veri      1.00
engine    0.98
feature   0.98
repress   0.98
extract   0.98
word      0.98
classify  0.97
financi   0.97
```

Figure 10. Identifying associated words

Term Frequency Calculation

Frequency of the terms taken from the ontology for every video is computed. Number of terms is calculated along with the semantic terms. If words appear frequently in a document, then they will be considered as important words. These words will be given a high score. But if a word appears in many documents, that word will get low score since it is not a unique identifier. Term frequency of documents are calculated using equation 9. Based on the frequency of terms and semantic terms (Figures 11-12), videos are ordered according to user query. Key frame selection and text extraction can be carried out at the initial stage itself even before the user starts querying. Association and semantic relationship meanings are found among the contents based on the user query information, so it can be carried out after the user gives the query.

	Terms								
Docs	recent	refer	relate	report	repres	resolv	result	science	sens
Text1.txt	4	9	2	2	1	3	2	2	2
Text2.txt	1	3	2	2	2	1	1	5	1
Text3.txt	2	1	3	1	1	4	1	4	1
	Terms								
Docs	site	social	special	support	system	technolog	time	train	univers
Text1.txt	1	1	4	3	13	1	8	8	4
Text2.txt	8	8	1	2	10	1	12	12	6
Text3.txt	6	5	1	6	8	3	1	1	6
	Terms								
Docs	updat	use	whole	word	work				
Text1.txt	1	2	3	17	18				
Text2.txt	4	13	1	4	1				
Text3.txt	1	36	1	7	1				

Figure 11. Document-wise term frequency

> freq

abstract	accord	all	analyz	base	better	build	case	center	confus
5	6	12	4	12	5	5	4	8	6
construct	content	data	develop	differ	each	effect	emot	example	exist
14	3	35	19	9	9	10	12	4	3
Find	first	focus	found	general	group	help	implement	import	improv
8	5	5	4	4	12	12	8	8	7
Inform	issu	journal	keyword	know	knowledg	main	make	mean	model
36	15	12	3	7	11	8	14	4	36
Network	object	one	onlin	other	out	own	paper	perform	practic
17	4	8	19	15	10	4	14	5	3
Present	prevent	problem	provid	public	purpos	recent	refer	relat	report
3	14	16	12	7	5	4	5	7	4
Repress	resolv	result	scienc	sens	site	social	special	support	system
14	4	8	10	3	15	15	6	11	31
Technolog	time	tool	train	univers	updat	use	whole	word	work
5	21	11	10	16	6	51	5	24	20

Figure 12. Overall term frequency

Grouping of Semantically Related Videos

Video database consists of video from four different domains (categories) as sample which includes agriculture, India, quantum optics and datamining. Domain ontology (Figure 13-14) is a formal model that serves as system's structure. Domain expert explores specific knowledge, analyse the most relevant entities and organises them into concepts and relationships. The skeleton of ontology consists of a hierarchy of generalized and specialized concepts. Domain ontology is created for each category in video database. Keywords of each video will be compared with the domain ontology. Videos are grouped on the basis of domain ontology and video keywords. When query matches with the cluster, all the videos in the cluster will be given as results.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:video="http://vit.ac.in#">
 <rdf:Description rdf:about="http://vit.ac.in#Video3">
  <video:course>DataMining Clustering Algorithms</video:course>
 </rdf:Description>
 <rdf:Description rdf:about="http://vit.ac.in#Video1">
  <video:course>DataMining Basics</video:course>
 </rdf:Description>
 <rdf:Description rdf:about="http://vit.ac.in#Video2">
  <video:course>DataMining Classification Algorithms</video:course>
 </rdf:Description>
</rdf:RDF>
```

Figure 13. Sample RDF used for retrieval using Jena

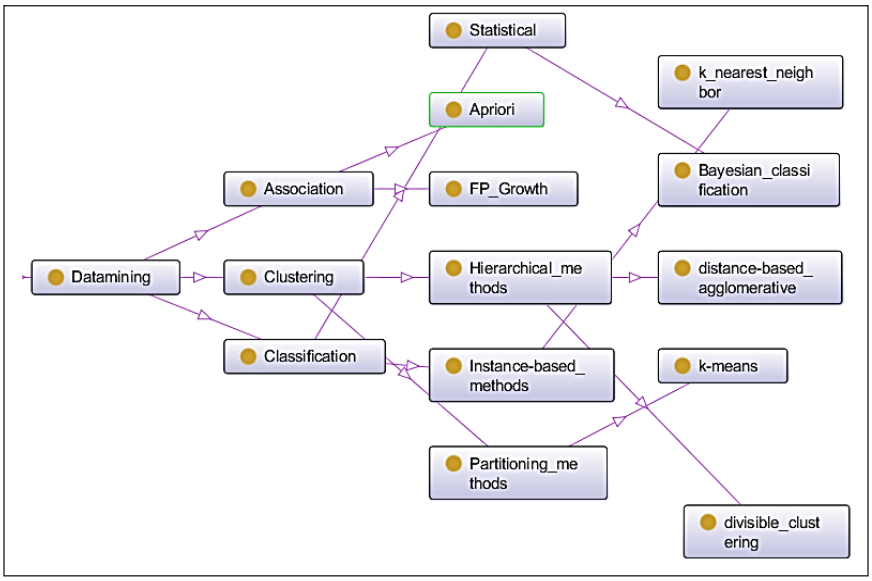


Figure 14. Graphical representation of relationship between entities

RESULTS AND DISCUSSION

Key frame extraction, text extraction, term frequency calculation and finding semantic meaning and relationships are the major components taken for result analysis. Key frames are selected randomly and from the key frames text are extracted through Tesseract, an open source API. WordNet, an open source electronic lexical database used for finding the semantic meaning and relationships of terms in document and in query. Term frequencies are calculated using Term Frequency and Inverse Document Frequency.

This research work includes various video samples from the fields of agriculture, India, quantum optics and datamining for the lecture video retrieval. Queries includes all the four categories of videos in the database. In total 50 video samples are taken for experimentation. Duration of each video is around one hour. Number of frames on each video depends on the quality of the video. Figure 15 explains the retrieved video contents for query. The performance of the proposed method is compared with the existing Correlation incorporated Naive Bayes (CNB) lecture video retrieval (Poornima & Saleena, 2018), since this is the most recent work done among all other related work. Ground truth was generated by experts who have a knowledge of the topic with an understanding of topics to follow. For a chosen subset of documents in these sets, ground truth was determined by generating potential results for sample documents that fall under the different categories of relatedness. To compare our results against the ground truth, we use precision (no. of relevant results/no. of results obtained), recall (no. of relevant results/no. of expected results) and f-measure metrics. Average of precision, recall and f-measures for queries of each category is calculated and included for those calculations.

Figure 15. Query and relevant videos retrieved

Precision is the ratio of the number of relevant videos retrieved to the total number of irrelevant and relevant videos retrieved. Figure (Figure 16) shows the precision rate of semantic information retrieval for query with CNB. Precision rate of semantic information retrieval achieves 0.9205 whereas precision rate for CNB is 0.72, which is low compared to semantic information retrieval. Thus, the values conclude that the precision rate is better for semantic information retrieval when compared to CNB.

Recall is the ratio of the number of relevant videos retrieved to the total number of relevant videos in the database. Figure (Figure 17) shows the recall rate of semantic information retrieval with CNB. Recall rate of semantic information retrieval achieves 0.9335 whereas recall rate for CNB method is 0.78, which is low compared to semantic information retrieval. Thus, the values conclude that the recall rate is better for semantic information retrieval when compared to CNB method. Comparison between precision and recall is shown in the figure (See Figure 18).

F-measure is the harmonic mean of precision and recall, that is, it is a combination of precision and recall. F-measure is intended to combine these two into a single measure of search effectiveness. Traditional equation for F-measure is given in equation 12:

$$F = 2.\frac{Precision.Recall}{Precision+Recall} \qquad [12]$$

Figure (Figure 19) shows the f-measure of semantic information retrieval for query with CNB. F-measure of semantic information retrieval achieves 0.925 whereas f-measure for CNB method is 0.75, which is low compared to semantic information retrieval. Thus, the values conclude that the f-measure is better for semantic information retrieval when compared to CNB.

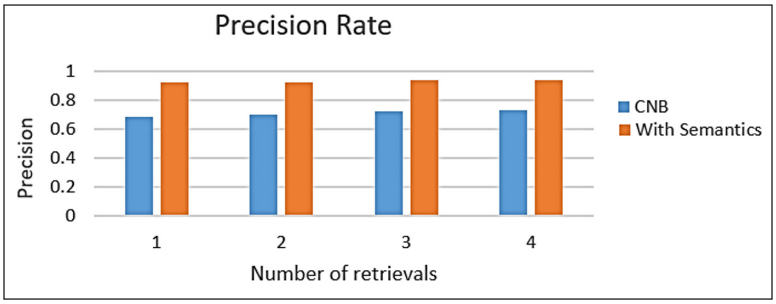


Figure 16. Precision rate

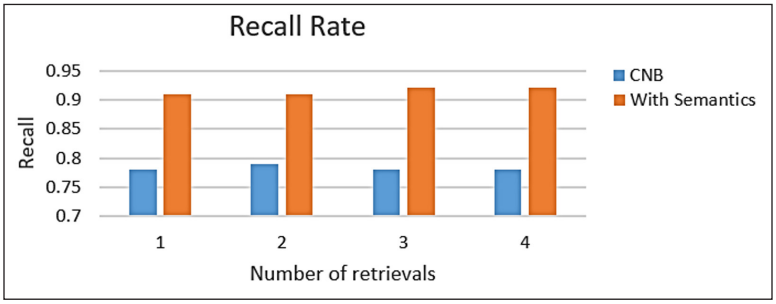


Figure 17. Recall rate

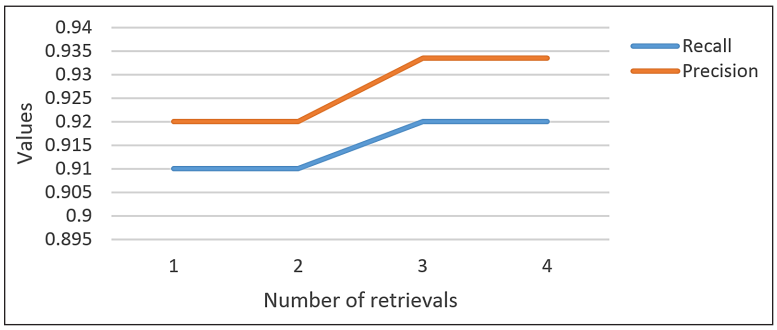


Figure 18. Precision and recall curve

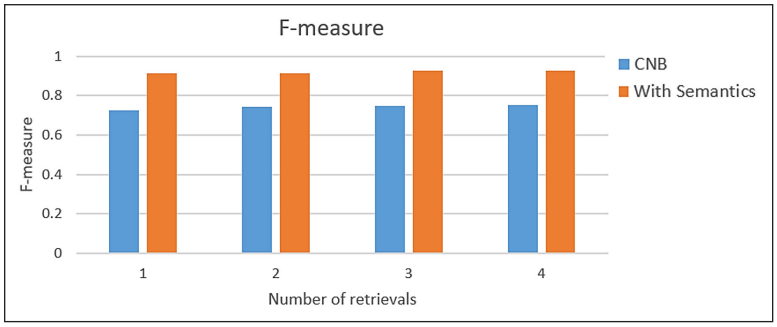


Figure 19. F-measure

CONCLUSION

Our proposed work has addressed the issues of extracting content descriptive annotations for the purpose of supporting content based video lecture retrieval. This paper has discussed a two phase methodology for capturing semantically related keywords from the video contents (visuals). First phase covers the extraction of text features from the video using connected components method and optical character recognition technique for keyframe detection and text extraction respectively. Second phase captures semantically related words for the frequently occurring words from the first phase. Videos are clustered by comparing semantically related words with domain ontology. The effectiveness of the proposed approach has been proven based on the experimentations done on that actual video set. Future work includes further experiments to extract keywords from all kinds of videos other than educational videos.

REFERENCES

Balasubramanian, V., Doraisamy, S. G., & Kanakarajan, N. K. (2016). A multimodal approach for extracting content descriptive metadata from lecture videos. *Journal of Intelligent Information Systems, 46*(1), 121-145.

Chatti, M. A., Marinov, M., Sabov, O., Laksono, R., Sofyan, Z., Yousef A. M. F., & Schroeder, U. (2016). Video annotation and analytics in CourseMapper. *Smart Learning Environments, Springer, 3*(10), 1-21.

Che, X., & Lin, L. (2015). A Survey of Current YouTube Video Characteristics. *IEEE Transaction on Multimedia, 22*(2), 56-63.

Cilibrasi, R. L., & Vitanyi, P. M. B. (2007). The Google Similarity Distance. *IEEE Transactions on Knowledge and Data Engineering, 19*(3), 370-383.

El-Khoury, V., Jergler, M., Bayou, G. A., Coquil, D., & Kosch, H. (2013). Fine-granularity semantic video annotation. *International Journal of Pervasive Computing and Communications, 9*(3), 243-269.

Ferman, A. M., & Tekalp, A. M. (2003). Two-stage hierarchical video summary extraction to match low- level user browsing preferences. *IEEE Transactions on Multimedia, 5*(2), 244-256.

Janwe, N. J., & Bhoyar, K. K., (2016). Video Key-Frame Extraction using Unsupervised Clustering and Mutual Comparison. *International Journal of Image Processing (IJIP), 10*(2), 73-84.

Jiang, Y. G., Dai, Q., Mei, T., Rui, Y., & Chang, S. F., (2015). Super Fast Event Recognition in Internet Videos. *IEEE Transactions on Multimedia, 17*(8), 1174-1186.

Kravvaris, D., Kermanindis, K. L., & Chorianopoulos, K. (2015, May). Ranking Educational Videos: The Impact of Social Presence. In *2015 IEEE 9th International Conference on Research Challenges in Information Science (RCIS)* (pp. 342-350). Athens, Greece.

Li, H., Liu, L., Sun, F., Bao, Y., & Liu, C. (2015). Multi-level feature representations for video semantic concept detection. *Neurocomputing, 172*, 64-70.

Li, K., Wang, J., Wang, H., & Dai, Q. (2015). Structuring Lecture Videos by Automatic Projection Screen Localization and Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(5), 1233-1246.

Muralikumar, J., Seelan, S. A., Vijayakumar, N., & Balasubramanian, V. (2016). A statistical approach for modeling inter-document semantic relationships in digital libraries. *Journal of Intelligent Information Systems, 48*(3), 477–498.

NPTEL Frequently Asked Questions. (2018, March 05). Retrieved July 8, 2017, from http://nptel.ac.in/faq.php

Poornima, N., & Saleena, B. (2018). Multi-modal features and correlation incorporated Naive Bayes classifier for a semantic-enriched lecture video retrieval system. *The Imaging Science Journal, 66*(5), 263-277.

Ravinder, M., & Venugopal, T. (2016). Content-Based Video Indexing and Retrieval using Key frames Texture, Edge and Motion Features. *International Journal of Current Engineering and Technology, 6*(2), 672-676.

Viana, P., & Pinto, J. P. (2017). A collaborative approach for semantic time-based video annotation using gamification. *Human-Centric Computing and Information Sciences, Springer, 7*(13), 1-21.

Wattanarachothai, W., & Patanukhom, K. (2015). Key Frame Extraction for Text Based Video Retrieval Using Maximally Stable Extremal Regions. In *2015 IEEE 1st International Conference on Industrial Networks and Intelligent Systems (INISCom)* (pp. 29-37). Tokyo, Japan.

Xu, Z., Mei, L., Liu, Y., Zhang, H., & Hu, C. (2015). Crowd Sensing Based Semantic Annotation of Surveillance Videos. *International Journal of Distributed Sensor Networks, 11*(6), 1-9.

Yang, H., & Meinel, C. (2014). Content Based Lecture Video Retrieval Using Speech and Video Text Information. *IEEE Transactions on Learning Technologies, 7*(2), 142-154.

Yu, X., Wang, L., Tian, Q., & Xue, P. (2004). Multilevel video representation with application to keyframe extraction. In *2004 IEEE 10th International Multimedia Modelling Conference* (pp. 117-123). Brisbane, Queensland, Australia.

Zhang, X., Liu, T., Lo, K., & Feng, J. (2003). Dynamic selection and effective compression of key frames for video abstraction. *Pattern Recognition Letters, 24*(9–10), 1523-1532.

Pertanika J. Sci. & Technol. 26 (4): 1591 - 1603 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Selection of the Best Location for a Small Hydro Power Project using the AHP-Weighted Sum and PROMETHEE Method

Jayantilal N. Patel* and Shilpesh C. Rana

Civil Engineering Department, Sarda Vallabhbhai National Institute of Technology, Surat, Gujarat, India

ABSTRACT

Hydro power is one of the renewable sources of energy which plays a significant role in the development of any country. Ranoli Branch Canal splits from Sakarda Branch Canal, one of the branch canals of Narmada Main Canal. The objective of this study is to find the best location for a small hydro power project out of four feasible locations on Ranoli Branch, using the Analytical Hierarchy Process (AHP)-weighted sum method. Result of weighted sum method has been validated using PROMETHEE method. The problem has been evaluated based on criteria for project cost, rated power, distance of the power house to the grid line, distance of the power house to the road, and distance of the power house to the village, and four canal fall locations, at chainage 7525 m, 9825 m, 17367 m, and 19844 m, as alternatives. The project cost was calculated by designing hydro power components (using Indian Standards Guideline) and applying actual market rates. The distance of the power house to the grid line, road, and village were obtained with the Google Play Store application called 'Map Distance Ruler Lite'. Optimisation resulted in the best location for hydro power generation in each canal. The fourth alternative, A_4 at chainage 19844 m, is the best location.

Keywords: AHP, branch canal, small hydro power project, weighted sum method

INTRODUCTION

The Sardar Sarovar project is one of the largest water resources projects of India. The Narmada Main Canal is a part of the Sardar Sarovar project, having 1133 m^3/s capacity at the head regulator. The length of the canal is 532 km. Narmada Main Canal has been divided in three phases. There are 25 branch canals, branching from the main canal phase-1 (Narmada Main Canal chainage-0 to144.50 km). Sakarda Branch Canal branches off from Narmada Main Canal at 102.95 km

ARTICLE INFO

Article history:
Received: 10 July 2017
Accepted: 03 April 2018
Published: 24 October 2018

E-mail addresses:
jnp@ced.svnit.ac.in (Jayantilal N. Patel)
shilpesh1978@yahoo.co.in (Shilpesh C. Rana)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

and is 35.13 km long. Ranoli Sub-Branch Canal branches off from Sakarda Branch Canal at 6.46 km downstream as shown in Figure 1.

On Ranoli Sub-Branch Canal, four falls have been identified as the locations to execute a small hydro power project, as shown in Figure 2. Out of four locations for a small hydro power project, the best location has been identified using AHP-weighted sum method. Four canal fall locations, at chainage 7525 m, 9825 m, 17367 m, and 19844 m, have been identified for the power house location. Five criteria, including project cost, rated power, distance of the power house from the grid line, distance of the power house from the road, and distance of the power house from the village, were considered in the selection of best the power house location.

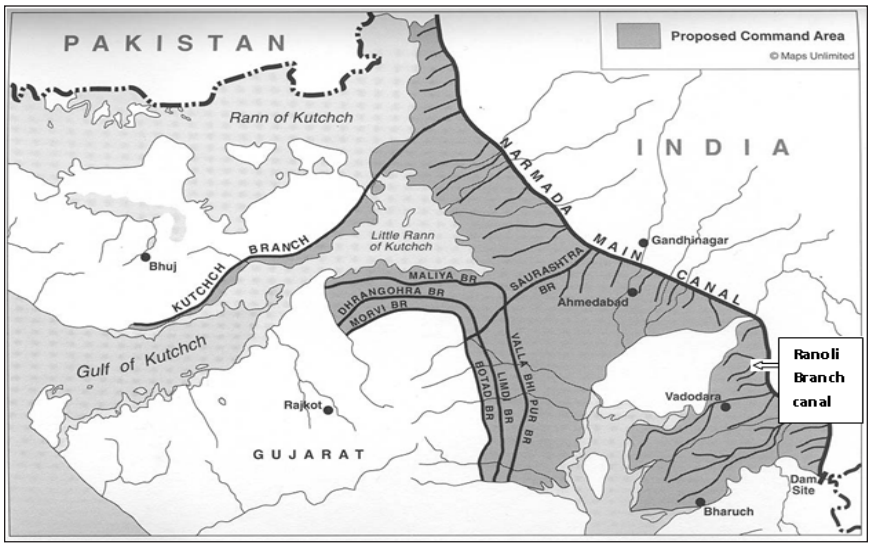


Figure1. Map of Narmada canal *(*Adapted from http://www.narmada.org/sardarsarovar.html*)*

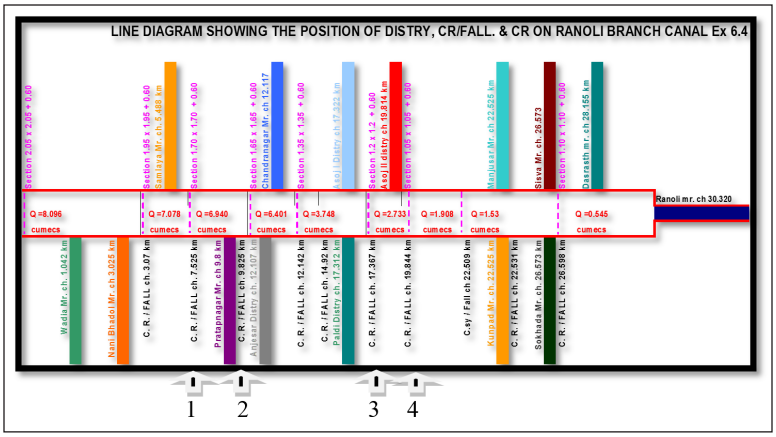


Figure 2. Proposed power house locations (adapted from Sardar Sarovar Narmada Nigam Limited, Kalol division)

METHODS

Analytical Hierarchy Process (AHP) method was used to select the best location for the power house. Calculations were based on expert opinion, project cost, rated power, distance of the power house from the grid line, distance of the power house from the road, and distance of the power house from the village. Canal falls at different chainages, as given in Table 1, were considered as alternatives.

Table 1
Canal fall locations

Fall No.	Chainage distance (m)
1	7525
2	9825
3	17367
4	19844

Analytical Hierarchy Process (AHP) Method

The analytical hierarchy process method was developed by Saaty (1977, 1980, 1982, and 1995) as a method for analysing decisions by structuring the decision's components (Bana & Vansnick, 2008; Turcksinaet al., 2011). AHP method was used to organise and analyse complex decisions (Saracoglu, 2013). The AHP method proves to be one of the most applicable methods of multi-criteria analysis (MCA) (Roman, 2012).

The method employs the following steps to solve a problem (Ioan et al., 2017):

(i) Identifying the problem.

(ii) Establishing the decision-making criteria. The criteria (objectives) are defined that shall be used for the selection of the alternatives. The data are written in a decision criteria matrix C = [C_J]. Here j = l...m represents the number of criteria (Naghiu et al., 2016).

(iii) Establishing the decision-making alternatives. In this stage, the set of alternatives that can be applied are identified, while the data are written in the alternatives matrix A = [Ai]. Here i = in, representing the number of alternatives (Naghiu et al., 2016).

(iv) Determining relative weight of criteria by comparing the criteria in pairs. In this step, the relative weight of the criteria c = [c_{ij}] is compared to their importance in making the decision (Prejmecrean, 2012).The relative weight of criteria is determined by performing pair-wise comparison.

(v) Normalising the comparisons between criteria. The normalised values "n_{ij}" are obtained by dividing the value found from comparison with the total value of their column (Dobrea, 2006) the calculation is based on

$$n_{ij} = \frac{Cij}{Sj} \quad [1]$$

Then, the pair-wise comparison between criteria is transformed into weights, which are calculated as an average of the normalised values on each row as:

$$kj = \frac{\sum_{j=1}^{m} nij}{m} \quad [2]$$

where kj represents the importance coefficients (weights) of the decision criteria. To use normalised values, the following condition must be satisfied:

$$\sum_{j=1}^{m} kj = 1 \quad [3]$$

(vi) Determining the consistency factor of the decision criteria matrix. To determine the consistency factor of the matrixes, the following steps are performed (Dobrea, 2006):

a) Determining the vector of priorities - λmax. The vector of priorities is calculated as an average of multiplication between the matrix of relative weights of decision criteria and the average weight of decision criteria as follows:

$$\lambda max = \sum_{j=1}^{m} \frac{(c\,x\,k)j}{m\,x\ kj} \quad [4]$$

where (c k) j represent the elements of the matrix vector determined because of multiplying the "c" matrix with "k" vector (Dobrea, 2006).

b) Determining the uniformity coefficient. The uniformity coefficient "CI" is calculated as:

$$CI = \frac{\lambda max - m}{m-1} \quad [5]$$

c) Determining the consistency factor of the matrixes. The consistency factor of matrixes "CR" is calculated as:

$$CR = \frac{CI}{R} \quad [6]$$

When determining the consistency relation, one considers the following rule: if CR< 0.10, then the matrix is consistent, meaning that the vector of the weights is well-determined.

Weighted Sum Model

In decision theory, the weighted sum model (WSM) is the best known and simplest multi-criteria decision analysis (MCDA) or method for evaluating a number of alternatives with a number of decision criteria. The WSM is probably the most commonly used approach,

especially in single dimension problems (Triantaphyllou & Mann, 1989). It is very important to state here that it is applicable only when all the data are expressed in exactly the same unit. If this is not the case, then the result is equivalent to "adding apples and oranges." Importance of alternative A_i, denoted as $A_i^{\text{WSM-score}}$, is defined as follows:

$$A_i^{\text{WSM-score}} = \sum_{n=1}^{\infty} (w_j a_{ij,}) \text{ for } i = 1, 2, 3, \ldots, m \text{ and } j=1, 2, 3, \ldots, n \quad [7]$$

For the maximisation case, the best alternative is the one that yields the maximum total performance value (Fishburn, 1967).

PROMETHEE Method

The PROMETHEE method is one of the MCDM methods that have been developed by Brans (Mohammad et al., 2017). It is a quite simple ranking method in conception and application compared with other methods for multi-criteria analysis. (Amir et al., 2007). The advantage of PROMETHEE is ranking and selecting among criteria which are conflicting (Mohammad et al., 2017). As per Ivica et al. (2015) an input for PROMETHEE method was a matrix consisting of set of potential alternatives (actions) A, where each a element of A has its $f_j(a)$ which represented evaluation of criteria j. Each evaluation fj(ai) must be a real number. Method PROMETHEE I rank actions by a partial pre-order, with the leaving flow:

$$\phi^+(a) = \frac{1}{n-1} \sum_{x \in A}^{q} \pi(a, x) \quad [8]$$

and entering flow:

$$\phi^-(a) = \frac{1}{n-1} \sum_{x \in A}^{q} \pi(x, a) \quad [9]$$

Where n denotes the number of actions, 'a' is a set of actions and π is the aggregated preference index defined for each couple of actions. The PROMETHEE I method gives partial pre-order. PROMETHEE II gives a net outranking flow which ranks the actions by total pre-order:

Ranking of actions is possible by determining the net outranking flow φ(a) (Tijana et al., 2017)

$$\phi(a) = \phi^+(a) - \phi^-(a) \quad [10]$$

Where $\phi^+(a)$ is leaving flow and $\phi^-(a)$ is entering flow.

ANALYSIS AND RESULTS

The best location for the power house has been found stepwise, as below using AHP, WSM and PROMETHEE.

Analytical Hierarchy Process (AHP) Method

(I) Identification of Problem. The problem is to find out the best location of the power house from four available alternatives.

(II) Criteria Consideration. Five criteria have been considered, as provided in Table 2.

Table 2
Set of decision criteria for selection of best location of a hydro power station

No.	Criteria	Name of criteria
1	C1	Project cost
2	C2	Rated power
3	C3	Distance of power house from grid line
4	C4	Distance of power house from road
5	C5	Distance of power house from village

(III) Determining alternatives. Canal falls located at various chainages have been considered as alternatives, as shown in Table 3.

Table 3
Set of alternatives for selection of the best location of a hydro-power station

No.	Fall location from the head regulator in meter	Alternative
1	7525	A1
2	9825	A2
3	17367	A3
4	19844	A4

(IV) Comparison of Criteria of the Pair to Determine Relative Weight of Criteria. Relative weights of five decision-making criteria have been compared to the next upper hierarchy rank, as given in Table 4. In the matrix, the diagonal values have been assigned as 1. The entire matrix has been prepared, considering that if C1 is 1.5 times more preferred than C2, and then criteria C2 is1/1.5 times less preferred than criteria C1.

Table 4
Values of relative weight for comparison between criteria (A_1 matrix)

	C1	C2	C3	C4	C5
C1	1	1.5	2	2.5	3
C2	1/1.5	1	2	3	3
C3	1/2.0	1/2.0	1	2	2
C4	1/2.5	1/3.0	1/2.0	1	1
C5	1/3.0	1/3.0	1/2.0	1	1

$$A_2 = \begin{bmatrix} 1.86/5.651 = 0.33 \\ 1.644/5.651 = 0.3 \\ 1.000/5.651 = 0.177 \\ 0.582/5.651 = 0.103 \\ 0.561/5.651 = 0.1 \end{bmatrix}$$

$A_3 = A_1 \text{ X } A_2$ [11]

$$A_3 = \begin{bmatrix} 1.6615 \\ 1.483 \\ 0.9 \\ 0.5235 \\ 0.5236 \end{bmatrix}$$

Relative normalised weights (w_j) of the attribute are calculated:

W_{C1}= 1.6615, W_{C2}=.300, W_{C3}=0.90, W_{C4}=0.5235, W_{C5}=0.5015

$A_4 = A_3 / A_2$ [12]

Avg. A_4 = 5.051

λ_{max} = Avg. A_4 [13]

= 5.051

Uniformity Co-efficient (C. I.) = (λ_{max}-M) / (M-1) [14]

= 0.013

Consistency Factor (C.R.) = C.I. /R.I. [15]

= 0.0112< 0.1

Weighted Sum Method

Using the relation $A_i^{\text{WSM-score}} = \sum_{n=1}^{\infty} (w_j a_{ij,})$ for i= 1, 2, 3…m. [16]
The scores of four alternatives were calculated.

Table 5
Normalised relative weights for decision criteria matrix

	C_1	C_2	C_3	C_4	C_5
Weightage	0.330	0.300	0.177	0.103	0.100
A_1	0.461	1.000	0.136	1.000	0.346
A_2	0.437	0.810	0.142	0.438	0.385
A_3	0.815	0.517	0.624	0.348	0.230
A_4	1.000	0.380	1.000	0.672	1.000

The WSM score for the four alternatives are:

$A_1^{\text{WSM-score}} = 0.461 \times 0.330 + 1 \times 0.3 + 0.136 \times 0.177 + 1 \times .103 + 0.346 \times .1 = 0.4769$

Similarly,

$A_2^{\text{WSM-score}} = 0.4959$, $A_3^{\text{WSM-score}} = 0.5934$ and $A_4^{\text{WSM-score}} = 0.7902$

PROMETHEE Method

The original evaluation matrix is shown in Table 6.

Table 6
The original evaluation matrix

Power House Location	Project Cost	Rated Power	Grid line distance from site	Road distance from site	Village distance fron site
Chainage	C_1	C_2	C_3	C_4	C_5
M	₹	kW	km	km	km
7525	8143274	230.0	3.900	1.790	1.590
9825	8592458	185.4	3.740	4.090	1.430
17367	4609490	119.0	0.850	5.142	2.410
19844	3755189	87.9	0.530	2.665	0.550

Pair wise difference between values of alternatives for each criteria (five) has been performed. For criteria C1 (Project cost), Pair wise difference between A1 and A2 are 8143274-8592458 =-449184. Similarly pair wise difference between alternatives A2 and A1 are 8592458-8143274= 449184. Pair wise difference between alternative A1 with A1 is zero. The matrix has been generated considering 0 values for negative differences 1 value for positive differences.

Pair wise difference matrix and preference function values for criteria C1 to C5 are shown follows.

$$a = \begin{bmatrix} - & 1 & 0 & 0 \\ 0 & - & 0 & 0 \\ 1 & 1 & - & 0 \\ 1 & 1 & 1 & - \end{bmatrix} \quad b = \begin{bmatrix} - & 1 & 1 & 1 \\ 0 & - & 1 & 1 \\ 0 & 0 & - & 1 \\ 0 & 0 & 0 & - \end{bmatrix} \quad c = \begin{bmatrix} - & 0 & 0 & 0 \\ 1 & - & 0 & 0 \\ 1 & - & 0 & 0 \\ 1 & 1 & 1 & - \end{bmatrix}$$

$$d = \begin{bmatrix} - & 1 & 1 & 1 \\ 0 & - & 1 & 0 \\ 0 & 0 & - & 0 \\ 0 & 1 & 1 & - \end{bmatrix} \quad e = \begin{bmatrix} - & 0 & 1 & 0 \\ 1 & - & 1 & 0 \\ 0 & 0 & - & 0 \\ 1 & 1 & 1 & - \end{bmatrix}$$

Multi-criteria preference Index π(a, x) for pair wise alternatives (A1,A2) is computed as follows: preference function values for A1 and A2 for criteria C1 to C5 are 1,1,0,1,0 and corresponding weights of criteria are 0.330, 0.300, 0.177, 0.103, 0.100

π(A1, A2)= .33X1+.30X1+.177X0 +.103X1+.1X0 = .733

Similarly,

A_1A_1= - , A_1A_2=0.733, A_1A_3=0.503, A_1A_4=0.403, A_2A_1=0.277, A_2A_2= - , A_2A_3=0.503, A_2A_4=0.3, A_3A_1=0.507, A_3A_2=0.507, A_3A_3= - , A_3A_4=0.3, A_4A_1=0.607, A_4A_2=0.71, A_4A_3=0.71, A_4A_4= -.

Table 7
Multi-criterion preference index values

Alternative	A1	A2	A3	A4
A1	0	0.733	0.503	0.403
A2	0.277	0	0.503	0.300
A3	0.507	0.507	0	0.300
A4	0.607	0.710	0.710	0

Computation of leaving, entering and net flow values

Table 8
Leaving, entering and net flow values for different alternatives

Alternative	$\phi^+(a)$	$\phi^-(a)$	$\phi(a)$	Rank
A1	0.546	0.464	0.0826	2
A2	0.360	0.650	-0.290	4
A3	0.438	0.572	-0.134	3
A4	0.676	0.334	0.342	1

Table 9
Normalised relative weights for decision criteria matrix

	C_1	C_2	C_3	C_4	C_5
Weightage	0.330	0.300	0.177	0.103	0.100
A_1	0.461	1.000	0.136	1.000	0.346
A_2	0.437	0.810	0.142	0.438	0.385
A_3	0.815	0.517	0.624	0.348	0.230
A_4	1.000	0.380	1.000	0.672	1.000

Table 10
Leaving, entering and net flow values for different alternatives

Alternative	$\phi^+(a)$	$\phi^-(a)$	$\phi(a)$	Rank
A1	0.546	0.464	0.0826	2
A2	0.360	0.650	-0.290	4
A3	0.438	0.572	-0.134	3
A4	0.676	0.334	0.342	1

Using the relation:

$A_i^{\text{WSM-score}} = \sum_{n=1}^{\infty} (w_j a_{ij,})$ for I = 1, 2, 3…m. [17],

the scores of four alternatives were calculated.

Table 11
Normalised relative weights for decision criteria matrix

	C_1	C_2	C_3	C_4	C_5
Weightage	0.330	0.300	0.177	0.103	0.100
A_1	0.461	1.000	0.136	1.000	0.346
A_2	0.437	0.810	0.142	0.438	0.385
A_3	0.815	0.517	0.624	0.348	0.230
A_4	1.000	0.380	1.000	0.672	1.000

The WSM score for the four alternatives are:

$A_1^{\text{WSM-score}} = 0.461 \times 0.330 + 1 \times 0.3 + 0.136 \times 0.177 + 1 \times .103 + 0.346 \times .1$

$= 0.4769$

Similarly,

$A_2^{\text{WSM-score}} = 0.4959$, $A_3^{\text{WSM-score}} = 0.5934$ and $A_4^{\text{WSM-score}} = 0.7902$

DISCUSSION

Present study was to select the best location of small hydro-power plant. Five criteria project cost, rated power, distance of the power house from the grid line, distance of the power house from the road, and distance of the power house from the village, were considered in the selection of best the power house location. Four canal fall locations, at chainage 7525 m, 9825 m, 17367 m, and 19844 m, had been identified for the power house location as an alternative.

Here the weightage were given to project cost (33%), followed by rated power (30%), distance of the power house from the grid line (17.7%), distance of the power house from the road (10.3%), distance of the power house from the village (10%). The AHP method provided a consistency factor of 0.0112.

The values of the WSM score, *net flow values* and ranks for the four alternatives are listed in Table 12.

Table 12
Values of WSM-Score, Net flow values and rank

Alternative	Fall location from the head regulator in meter	WSM		PROMETHEE	
		Score	Rank	Net flow values	Rank
A1	7525	0.4769	4	0.0826	2
A2	9825	0.4959	3	-0.290	4
A3	17367	0.5934	2	-0.134	3
A4	19844	0.7902	1	0.342	1

As a result of WSM and PROMETHEE alternative 4 is the best alternative for the location of small hydro power project.

CONCLUSION

As a result of AHP method the Consistency Factor (C.R.) is 0.0112 which is less then 0.10, which indicates that the decision criteria matrix is consistent.

Weighted Sum and PROMETHEE method were used to find the best location of small hydro power project out of four alternatives (Canal fall locations at chainage 7525 m, at chainage 9825 m, at chainage 17367 m and at chainage 19844 m).

From the results of both Weighted Sum and PROMETHEE methods the best location for small hydro power project is at 19844 m Chainage (A4 alternative).

REFERENCES

Amir, A., Chaharsooghi, K., & Akbar, E. (2007). Decision making in stock trading: An application of PROMETHEE. *European Journal of Operation Research, 177*(2), 673-683.

Bana, C., & Vansnick, J. (2008). A critical analysis of the Eigen value method used to derive priorities in AHP. *European Journal of Operational Research, 187*(3), 1422-1428.

Anjasmoro, B., & Sangkawati, S. (2017). Priority Analysis of Small Dams Construction using Cluster Analysis, AHP and Weighted Average Method Case Study: Small Dams in Semarang District. *Procedia Engineering, 171*, 1514-1525.

Dobrea, R. (2006). *Eficienta modernizarii sistemelor tehnico-economice [Efficiency of modernization of technical and economical systems]* (Doctoral dissertation), The Bucharest University of Economic Studies, Bucharest, Romania.

Fishburn, P. *(1967). Additive Utilities with Incomplete Product Set: Applications to Priorities and Assignments. Operations Research, 15(3), 537-542.*

Ioan, A., Gheorghe, B., Ioan, G., George, N., & Florin, G. (2017). Choosing the optimal technology to rehabilitate the pipes in water distribution systems using the AHP method. *Energy Procedia, 112*, 19-26.

Ivica, V., Stipo, C., & Ivan, P. (2015). Competences-based Comparison and Ranking of Industrial Enterprised using PROMETHEE Method. *Procedia Engineering, 100*, 445-449.

Nikouei, M. A., Oroujzadeh, M., & Mehdipour-Ataei, S. (2017). The PROMETHEE multiple criteria decision making analysis for selecting the best membrane prepared from sulfonated poly (ether ketone) s and poly (ether sulfone) s for proton exchange membrane fuel cell. *Energy, 119*, 77-85.

Naghiu, G., Giurca, I., Achilean, I., & Badea, G. (2016). Multicriterial analysis on selecting solar radiation concentration for photovoltaic panels using Electre-Boldur method. *Procedia Technology, 22*, 773-780.

Prejmecrean, V. (2012). Sisteme pentru fundamentarea deciziilor (Systems for Substantiating Decisions) [internet]. Cluj-Napoca: Universitatea Babes Bolyai: 2012. Cursul 5 Teoria Deeiziilor (Course 5 Decisions). Retrieved June 17, 2015, from http://www.cs.ubbcluj.ro/—per/Dss/Dss_5.pdf. p. 13.

Roman, M. (2012). *Analiza multi criteriala manual* [Handbook – Multi-criteria Analysis]. Bucuresti, Romania: Acadimia de Studie Economice.

Saracoglu, B. (2013). Selecting industrial investment locations in master plans of countries. *European Journal of Industrial Engineering, 7*(4), 416–441.

Tijana, V., & Nada, D. (2017). Multi-criteria decision analysis for sub- watersheds ranking via the PROMETHEE method*. International Soil and Water Conservation Research, 5*(1), 50-55

Triantaphyllou, E., & Mann, S. (1989). An examination of effectiveness of Multi-Dimensional Decision Making Methods: A Decision-Making Paradox. *Decision Support System, 5*(3), 303-312.

Turcksina, L., Bernardinia, A., & Macharisa, C. (2011). A combined AHP-PROMETHEE approach for selecting the most appropriate policy scenario to stimulate a clean vehicle fleet. *Proedia Social and Behavioral Sciences, 20*, 954-965.

Yoon, K. (1987). Reconciliation among discrete compromise situations. *Journal of Operational Research Society, 38*(3), 277–286.

Tijana, V., & Nada, D. (2017). Multi-criteria decision analysis for sub- watersheds ranking via the PROMETHEE method. *International Soil and Water Conservation Research, 5*(1), 50-55.

Pertanika J. Sci. & Technol. 26 (4): 1605 - 1618 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Evaluation of DSSAT- Ceres Model for Maize under Different Water and Nitrogen Levels

Abedinpour, M.[1]* and Sarangi, A.[2]

[1]*Water Engineering Department, Kashmar Higher Education Institute, Kashmar, Iran*
[2]*Water Technology Center, IARI, New Delhi, India*

ABSTRACT

Crop models can accurately estimate crop growth, biomass yield (BY) and grain yield (GY) with a priori information of the crop, soil properties and water management. Generation of new knowledge through traditional agricultural practices is not possible to meet the requirements for novel agro-technologies and they are generally season specific, expensive and time consuming. Hence, the CERES (Crop Environmental Resource Synthesis) model was calibrated using the data of 2009 and validated with the data of 2010 acquired from the field data of WTC, IARI, India. Irrigation applications comprised rainfed, i.e. no irrigation (I_1), irrigation at 50% of field capacity (FC) (I_2), at 75 % FC (I_3) and 100% FC or full irrigation (I_4). Nitrogen levels were: no nitrogen (N_1), 75 kg ha^{-1} (N_2) and 150 kg ha^{-1} (N_3). Model performance statistics of model efficiency (E), root mean square error (RMSE) and normalized root mean square error (NRMSE) were applied to evaluate the model performance. Model calibration for simulation of GY and BY provided prediction error statistics of $0.78<E<0.84$, $0.238<RMSE<0.70$ t ha^{-1} and $6<NRMSE<7$ %, respectively for all irrigation levels. Also, the model was validated for simulation of GY and BY for all treatment levels with the prediction error statistics of $0.86<E<0.88$, $0.36<RMSE<0.86$ t ha^{-1}, $0.95<R^2<0.98$ and $6<NRMSE<8$%. Nonetheless, it was observed that the CERES-maize model could be applied to estimate yield and biomass under the regional situations with reasonable accuracy.

Keywords: Calibration, maize, nitrogen, validation

ARTICLE INFO

Article history:
Received: 27 July 2017
Accepted: 26 April 2018
Published: 24 October 2018

E-mail addresses:
abedinpour_meysam@yahoo.com (Abedinpour, M.)
arjamadutta.sarangi@mail.mcgill.ca (Sarangi, A.)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Maize is the third most important grain crops after wheat and rice. According to advance prediction it is cultivated in 8.7 M ha, which covers 80% of cultivated area, in India. Maize grain production is more sensitive to lack or excess amounts of water and nitrogen fertilizer compared to the other cereals. Therefore, research on water and nitrogen management for enhancing maize productivity and use of appropriate crop model to simulate maize growth and yield assumes importance.

Crop modeling approaches a new possibility to educators, planners and policy makers to explore cultivar potential for new regions before conducting costly and time consuming field studies (Abedinpour et al., 2014). Today more than ever, raised crop production depends on judicious use of resources. In addition, issues such as climate change, soil carbon sequestration, long-term food security, and environmental sustainability have become important issues. Crop simulation models incorporating water, soil, plant and environment system can make a precious contribution to both furthering our understanding of the processes that determine crop responses and estimating crop performance, resource use, and environmental effects for different environments and management scenarios. The decision support system for agro-technology transfer (DSSAT) version 4.6 is a Windows-based computer program that comprises crop simulation models for over 42 crops. The model was established by database management programmers for soil, weather, and crop management and experimental data, and by utilities and application programs (Hoogenboom et al., 2010).

The CERES–maize model, which is a component of the Decision Support System for Agro- technological Transfer (DSSAT) is supported by data base management programs for soil, weather, and crop management and experimental data, and by utilities and application programs. The crop simulation models simulate growth, development and yield as a function of the soil-plant-atmosphere dynamics (Hoogenboom et al., 2017). The model has a capacity to simulate the daily crop growth, development and yield for variable soil and climatic conditions with various agronomic managements (Khaliq et al., 2007). The CERES-Maize response regarding yield simulation has been tested in Virginia (Hodges et al., 1987), and Australia (Hargreaves and McCown, 1988). CERES-Maize model simulates grain yield under water limiting conditions by calculating potential evaporation; and potential soil-water evaporation and potential plant-water transpiration are derived from potential evaporation and leaf area index. Simulations of deficit irrigation practices using models like the CSM-CERES-Maize can be used to look at numerous weather years and geographic locations. Amaral et al. (2015) indicated that, the CERES-maize model simulated maize growth, development and yield for both mineral fertilizer and poultry litter sources of nitrogen. Jianmei et al. (2014) evaluated the CERES model for wheat crop in Guanzhong Plain of Northwest China under different irrigation and nitrogen levels. The

results showed that the deviations of simulated BY, GY, leaf area index (LAI), cumulative evapotranspiration (ET) and crop water productivity (WP) from the observed values were reasonable, with NRMSE less than 21 %.

The aim of the research was to assess the performance of CERES model in simulating the impact of water and nitrogen fertilizer managements on growth and yield of maize in a semi-arid environment.

METHODS

Experimental Procedure

This study compares results from the CERES-maize model with observed data from a field experiment under rainfed, deficit and full irrigation in interaction with nitrogen levels at Water Technology Center (WTC) in Indian Agricultural Research Institute, (77° 8′ 45" to 77° 10′ 24" E longitude and 37° 22′ to 38° 39′ N latitude). The Meteorological Station is situated at 350m from the research field. The DSSAT model requires six weather parameters, including: daily minimum and maximum air temperature, daily relative humidity (%), rainfall, wind speed and, solar radiation. The collected weather parameters are presented in Figures 1and 2. The experiment was carried out using split-plot design based on randomized complete block design (RCBD) with three replications. Main factors were assigned with four irrigation levels *viz.* rainfed (I_1) and three irrigations at 50% (I_2), 75% (I_3) and 100% (I_4) of field capacity and three nitrogen levels [*viz.* non-fertilized (N_1), 75 (N_2) and 150 kg ha^{-1} (N_3)] as sub factors. Each plot consisted of 5 furrows spaced 0.75 m apart, with a furrow length of 4 m. Soil characteristics are given in Table 1. Maize cultivar BIO-9681 seed was sown at depth of 3 to 5 cm.

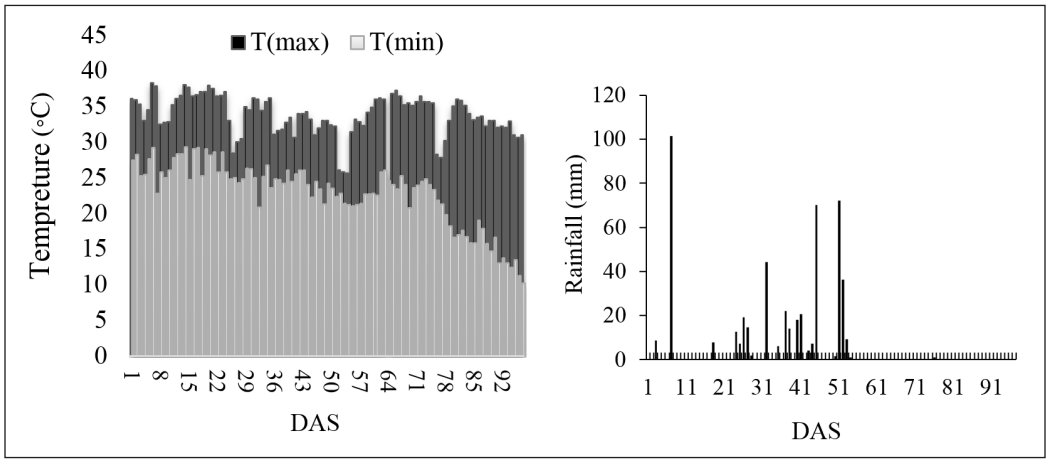


Figure 1. Daily maximum and minimum temperatures, and daily total rainfall during the crop growing season in 2009

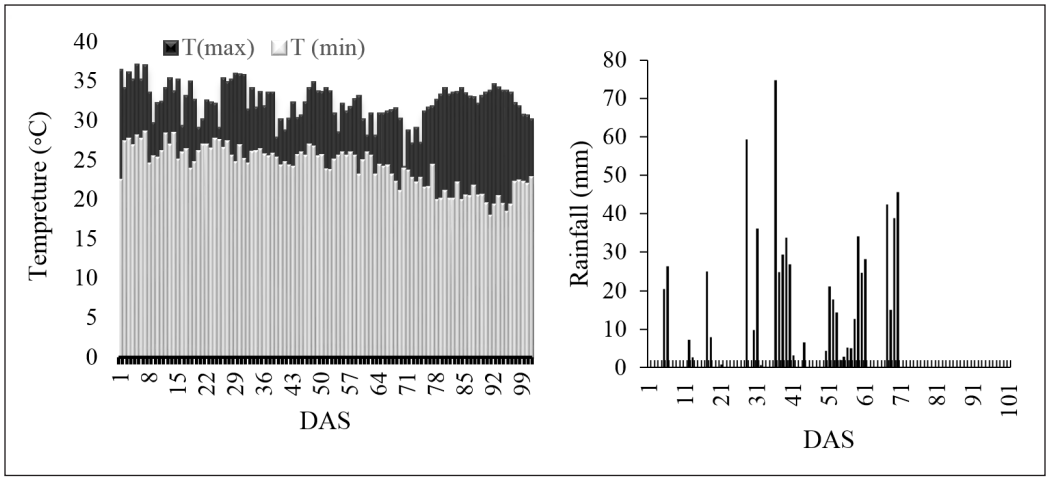


Figure 2. Daily maximum and minimum temperatures, and daily total rainfall during the crop growing season in 2010

Table 1
Physical and chemical properties of the soil of experimental field

Determination	Soil Depth (cm)				
	0-15	15-30	30-45	45-75	75-105
Sand (%)	52.4	53.7	44	39	38
Silt (%)	21	19	23	25	27
Clay (%)	26.6	27.3	33	36	35
Soil Texture	Sandy loam	Sandy loam	Loam	Loam	Clay loam
FC (w/w)	21.3	25.6	27.9	32.8	33.0
PWP(w/w)	9.5	10.2	13.7	14.7	15.0
K_S (cm day^{-1})	27.4	26.2	18.6	19.1	19.5
Bd (g cm^{-3})	1.41	1.43	1.39	1.37	1.36

Note: Bd: Bulk Density, Ks: Saturated Hydraulic Conductivity, FC: Field Capacity, PWP: Permanent Wilting Point

Assuming an effective root zone of 1.0 m, the total soil water content (SWC) in the top 1.0 m of soil was used in the analyses. One-third of total maize nitrogen requirement (N) was applied as basal dose. Additional N was applied with two split doses with one-third given at 21 days after sowing (DAS) and the remaining at 42 DAS of the crop. Measured quantity of irrigation water based on soil moisture content was directly applied to the furrows in experimental plots using High Density Poly Ethylene (HDPE) pipes. The conveyance loss was avoided by the use of HDPE pipes for supply of water from the source to all the experimental plots. In full irrigation treatment, water was applied up to field capacity level when soil moisture in the root zone approached 50% of total available water (TAW). In the deficit irrigation treatments (*i.e.* 50 and 75% of full irrigation), water was applied on the same day as the fully irrigated plot, but the irrigation depths were reduced

to 50 and 75% of the full irrigation treatment. There was no irrigation in the rainfed plots of the experiment. Irrigation water depths indicated by soil moisture deficit (SMD) in each treatment was calculated using soil-moisture content before irrigation, root zone depth of the plants and bulk density using Eq. 1.

$$SMD = \left(\theta_{fc} - \theta_i\right) \times \rho_b \times D_r \times f \qquad [1]$$

In Eq. (1),

SMD: Soil moisture deficit (mm), θ_{fc}: soil moisture at field capacity, θ_i: soil moisture before irrigation (weight basis in %), D_{rz}: depth of effective root zone (mm), ρ_b: bulk density of the given soil layer (g cm^{-3}) and f: coefficient of each irrigation treatment (i.e. 0, 0.5, 0.75 and 1).

Canopy development was measured in terms of growth stages, leaf area, root length, and above ground biomass on bi-weekly basis by removing two plants per plot. Date of emergence, maximum canopy cover (CC), duration of flowering, start of senescence, and maturity were also recorded. In each crop growth stages, green leaves were separated and leaf area of each plant measured by leaf area meter to obtain leaf area index (LAI). The LAI was converted to crop canopy cover (CC). Dry biomass of above ground plant at each crop growth stages were obtained by weighing it after keeping in an oven for 48 hours at 65^{o} C. Grain yield was measured as weight of harvested grain with 13% grain moisture. Total biomass yield was determined by taking the weight of above ground plant parts, including the grain.

Model Description

The Decision Support System for Agrotechnology Transfer, DSSAT, Version is a software application program that comprises crop simulation models for over 42 crops (as of Version 4.6). For DSSAT to be functional, it is supported by data base management programs for soil, weather, and crop management and experimental data, and by utilities and application programs. The crop simulation models simulate growth, development and yield as a function of the soil-plant-atmosphere dynamics. Also, DSSAT and its crop simulation models have been applied for many applications, ranging from on-farm and precision management to regional assessments of the impact of climate variability and climate change. The crop models require daily weather data, soil surface and profile information, detailed crop management and crop genetic information, and cultivar or variety information as input data. Crop model evaluation is accomplished by inputting the user's minimum data, running the model, and comparing outputs with observed data. By simulating probable outcomes of crop management strategies, DSSAT offers users information with which to rapidly appraise new crops, products, and practices for adoption.

Statistics for Model Evaluation

The prediction error (P_e), coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE) and model efficiency (E) were used as the error statistics to testing the calibration and validation outputs of the model. Model performance was tested using the following statistical parameters:

$$P_e = \frac{(S_i - O_i)}{O_i} \times 100 \quad [2]$$

$$E = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \quad [3]$$

where S_i and O_i are simulated and actual (observed) data, $\bar{O}$ is mean value of O_i and n is the number of observations.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S_i - O_i)^2}{n}} \quad [4]$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|S_i - O_i| \quad [5]$$

$$NRMSE = \frac{RMSE}{\overline{O}} \times 100 \quad [6]$$

Model efficiency (E) and R^2 approaching one, and Pe, MAE, NRMSE and RMSE close to zero were indicators for better model performance. At last, the simulation is done significantly well with a NRMSE < 10%, good if 20 > NRMSE > 10, fair if 30 > NRMSE > 20, and poor if NRMSE >30% (Jamieson et al., 1991).

RESULTS

Grain yield, Biomass and Water Productivity

Grain yield, above ground biomass, and water productivity (WP) under non limiting fertilized (N_3), moderately fertilized (N_2) and poorly fertilized (N_1) conditions for 2009 and 2010 experiments are listed in Table 2. The lowest grain and biomass yields were observed to be 1430 and 6430 kg ha^{-1}, respectively in rainfed (I_1) and non-fertilized (N_1) treatment

and the highest yields were 5930 and 18150 kg ha^{-1}, respectively, under full irrigation (I_4) and recommended dose of nitrogen (N_3). The water productivity ranged from a minimum of 5.7 kg ha^{-1} mm^{-1} to a maximum of 12.9 kg ha^{-1} mm^{-1} in 2009. Water productivity for full irrigation (I_4) under N_3 treatment was the highest, whereas that for rainfed (I_1) treatment under non fertilized (N_1) condition was the lowest. During 2010, the highest (i.e. 12.4 kg ha^{-1} mm^{-1}) and lowest (6.9 kg ha^{-1}mm^{-1}) water productivity were obtained for I_1N_3 and both for I_4N_1 and I_3N_1 treatment combinations, respectively.

Table 2
Crop water use, irrigation water depths, grain yield, above ground biomass, water productivity (WP) and irrigation water use efficiency (IWUE) under varying N-fertilizer levels during 2009 and 2010

Year	Treatment	IWA (mm)	CWU (mm)	GY (kg ha^{-1})	WP (kg ha^{-1} mm^{-1})	IWUE (kg ha^{-1} mm^{-1})	Biomass (kg ha^{-1})
Non-limiting fertilizer level (N_3)							
2009	I_1	0	250	2360	9.4	NA	10240
	I_2	105	355	3625	10.2	12.56	14010
	I_3	158	408	4250	10.4	10.25	14670
	I_4	210	460	5930	12.9	15.7	18140
2010	I_1	0	423	5250	12.4	NA	16430
	I_2	24	447	5422	12.1	7.16	16370
	I_3	39	462	5525	11.9	7.05	17370
	I_4	58	481	5775	12.0	9.05	17600
Moderate-limiting fertilizer level (N_2)							
2009	**I_1**	**0**	**250**	**1950**	**7.8**	**NA**	**7950**
	I_2	**105**	**355**	**3190**	**9.0**	**11.81**	**10540**
	I_3	**158**	**408**	**4450**	**10.2**	**13.92**	**12390**
	I_4	**210**	**460**	**5120**	**11.1**	**15.1**	**14900**
2010	I_1	0	423	4535	10.7	NA	14100
	I_2	24	447	4685	10.5	6.25	14230
	I_3	39	462	4815	10.4	7.17	14620
	I_4	58	481	4785	9.9	4.31	14650
Poor fertilizer level (N_1)							
2009	**I_1**	**0**	**250**	**1430**	**5.7**	**NA**	**6400**
	I_2	**105**	**355**	**2535**	**7.1**	**10.52**	**8950**
	I_3	**158**	**408**	**3015**	**7.39**	**10.03**	**9360**
	I_4	**210**	**460**	**3395**	**7.38**	**9.35**	**10420**
2010	I_1	0	423	3160	7.5	NA	10170
	I_2	24	447	3245	7.3	3.54	10100
	I_3	39	462	3180	6.9	5.1	10200
	I_4	58	481	3315	6.9	2.67	10390

Note: CWU: Crop water used; IWA: Irrigation water applied; GY: Grain yield

Calibration of CERES-maize Model

Ceres model was calibrated using experimental data of 2009 to predict grain and biomass yields under different water and fertilizer application levels in the experiment. It was observed that, the maximum and minimum errors in grain yield prediction were in I_1N_1 and I_4N_3 treatments accounting 17% and 3%, respectively. The prediction errors in biomass for I_1N_1 and I_4N_3 treatments were 10% and 3%, respectively (Tables 3, 4 and 5). The model was calibrated for simulation of yield and biomass for all treatment levels with the prediction error statistics $0.78<E<0.84$, $0.238<RMSE<0.701$ t ha^{-1} and $6<NRMSE<7\%$ in simulating the yield and biomass for all irrigation levels. The result of model evaluation is presented in Table 6. The CERES model was able to predict the grain yield with good accuracy. The predicted biomass yield is illustrated in Figure 4**.**

Table 3
Calibrated values of above ground biomass, and grain yield of maize under different irrigation water regimes and non-limiting fertilizer doses (N_3)

Treatment	Yield (t ha^{-1})		Error (±%)	Biomass (t ha^{-1})		Error (±%)
	Measured	Simulated		Measured	Simulated	
Rain fed	2.36	2	-10	10.24	9.35	-9
I_2	3.62	3.38	-7	14.01	13.21	-6
I_3	4.25	4.48	5	14.67	13.95	-5
I_4	5.93	5.75	-3	18.14	17.53	-3

Table 4
Calibrated values of above ground biomass, and grain yield of maize under different irrigation water regimes and moderate-limiting fertilizer doses (N_2)

Treatment	Yield (t ha^{-1})		Error (±%)	Biomass (t ha^{-1})		Error (±%)
	Measured	Simulated		Measured	Simulated	
Rain fed	1.95	1.80	-7	7.95	7.34	-8
I_2	3.19	2.84	-11	10.54	9.93	-6
I_3	3.34	3.05	-9	12.39	11.37	-8
I_4	5.12	4.71	-8	14.9	13.50	-9

Table 5
Calibrated values of above ground biomass, and grain yield of maize under different irrigation water regimes and full-limiting fertilizer doses (N_1)

Treatment	Yield (t ha^{-1})		Error (±%)	Biomass (t ha^{-1})		Error (±%)
	Measured	Simulated		Measured	Simulated	
Rain fed	1.43	1.18	-17	6.4	5.74	-10
I_2	2.54	2.24	-12	8.95	8.26	-8
I_3	3.01	2.67	-11	9.36	8.65	-8
I_4	3.65	3.39	-7	10.42	9.77	-6

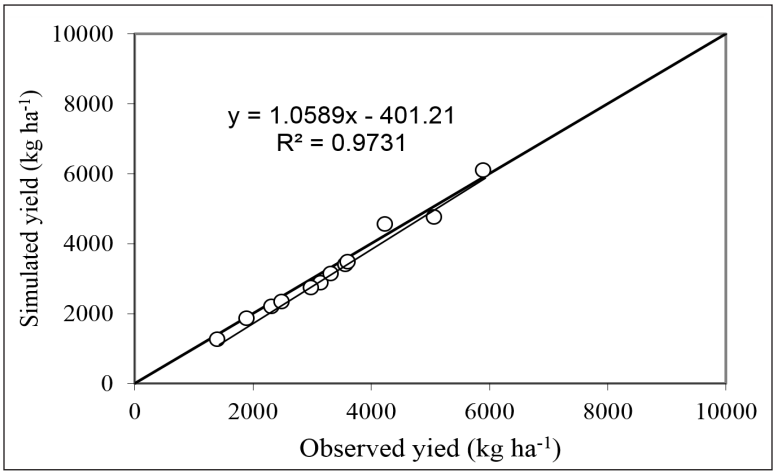


Figure 3. Simulated versus observed grain yield under all treatments

Table 6
Model calibration statistics for grain and biomass yields of maize

Crop parameters (t ha^{-1})	X_{obs}	X_{sim}	R^2	E	P(t)	RMSE	NRMSE (%)
Grain	3.36	3.16	0.97	0.84	0.47	0.238	7
Biomass	11.5	10.80	0.98	0.78	0.14	0.701	6

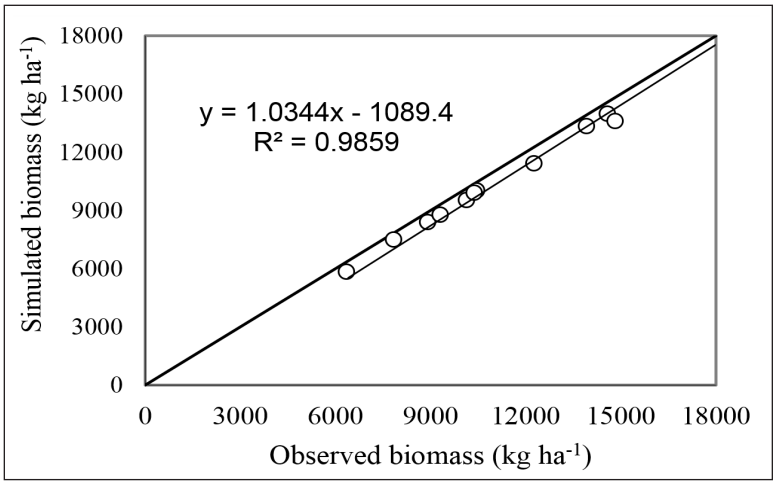


Figure 4. Simulated versus observed biomass yield for all treatments

Validation of Ceres model

Ceres model was validated using experimental data of 2010 to predict grain yield and biomass under different water and fertilizer application levels. It was observed that the maximum and minimum errors of grain yield prediction was obtained in I_1N_1 and I_4N_3

treatments at the rate of 15% and 3%, respectively. Similar this prediction was observed for biomass in I_1N_1 and I_4N_3 treatments by 11% and 2.9%, respectively (Tables 7,8 and 9). The model was validated for simulation of yield and biomass for all treatment levels with The prediction error statistics were $0.86<E<0.88$, $0.36<RMSE<0.86$ t ha^{-1}, $0.95<R^2<0.98$ and $6<NRMSE<8$ % in simulating the yield and biomass for all irrigation levels. The simulated versus observed values are illustrated in Figure 5 for grain yield and in figure 6 for biomass.

Table 7
Validation results of above ground biomass, and grain yield of maize under different irrigation water regimes and non-limiting fertilizer doses (N3)

Treatment	Yield (t ha^{-1})		Error (±%)	Biomass (t ha^{-1})		Error (±%)
	Measured	Simulated		Measured	Simulated	
Rain fed	5.12	4.75	-7	16.43	15.24	-7
I_2	5.62	5.30	-6	16.37	15.48	-5
I_3	5.52	5.78	5	17.37	17.81	3
I_4	5.77	5.58	-3	18.61	18.07	2.9

Table 8
Validation results of above ground biomass, and grain yield of maize under different irrigation water regimes and moderate-limiting fertilizer doses (N_2)

Treatment	Yield (t ha^{-1})		Error	Biomass (t ha^{-1})		Error
	Measured	Simulated	(±%)	Measured	Simulated	(±%)
Rain fed	4.53	4.05	-8	14.10	13.00	-8
I_2	4.68	4.12	-14	14.23	13.17	-7
I_3	4.80	4.21	-12	14.62	13.34	-9
I_4	4.87	4.35	-11	14.65	13.48	-8

Table 9
Validation results of above ground biomass, and grain yield of maize under different irrigation water regimes and full-limiting fertilizer doses (N_1)

Treatment	Yield (t ha^{-1})		Error	Biomass (t ha^{-1})		Error
	Measured	Simulated	(±%)	Measured	Simulated	(±%)
Rain fed	3.16	2.70	-15	10.17	9.02	-11
I_2	3.24	2.85	-12	10.20	9.19	-10
I_3	3.18	2.80	-12	10.25	9.34	-9
I_4	3.31	2.98	-10	10.39	9.57	-8

Figure 5 shows the performance of the model in terms of observed versus simulated grain yield. The regression line was more or less near to 1:1 line, indicating that the model was performing well under the test of different water and nitrogen levels. Similarly,

goodness of fit (R^2) as well as regression coefficients between observed and simulated data was significant. The coefficient of prediction was to the extent of 95% in case of trend run between the observed and simulated values.

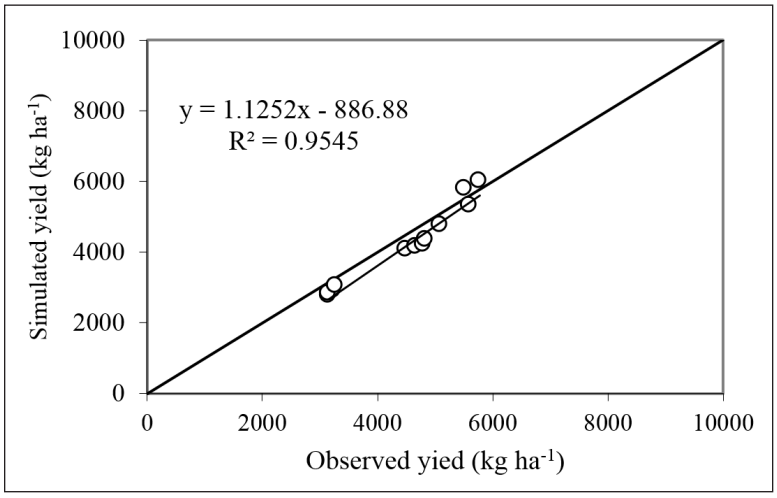


Figure 5. Simulated versus observed grain yield under all treatments

Table 10
Model validation statistics for grain and biomass yields of maize

Crop parameters (t ha^{-1})	X_{obs}	X_{sim}	R^2	E	P(t)	RMSE	NRMSE (%)
Grain	4.48	4.158	0.95	0.86	0.24	0.36	8
Biomass	13.86	13.06	0.98	0.88	0.027	0.86	6

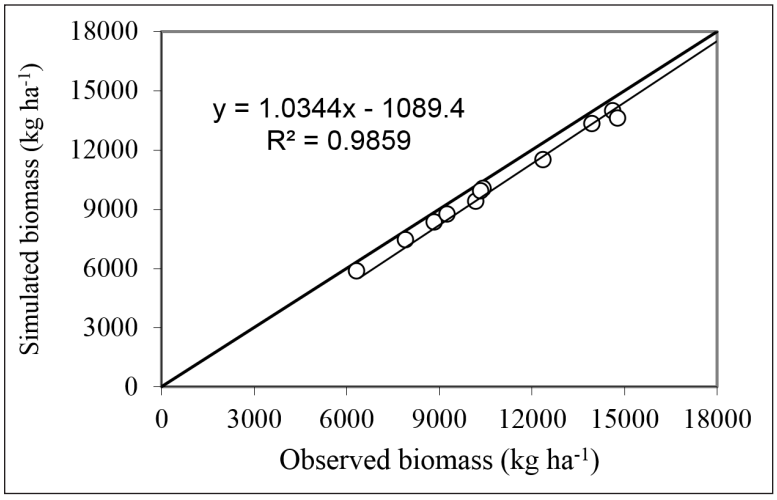


Figure 6. Simulated versus observed biomass yield under all treatments

Also, figure 6 shows the performance of the model in terms of observed versus simulated biomass yield. Goodness of fit (R^2) as well as regression coefficients between observed and simulated data was significant. The coefficient of prediction was to the extent of 98% in case of trend run between the observed and simulated values. Similarly, the regression line was near to 1:1 line, indicating that the model was performing well for maize crop under the test of different water and nitrogen levels.

DISCUSSION

Although in general, the grain and biomass yields were simulated by CERES-maize model correctly. However, in some case studies, the model had a slight trend of underestimating low observed yields. The result of this study is in agreement with findings by Panda et al. (2004) and Ló Pez-Cedrón et al. (2008). Also, a report showed that simulated mean grain yield was within 5% of measured grain yields for nine locations in the United States. But, Dogan et al. (2006) reported the opposite trend; however, this study had very poor correlation between simulated and observed yield values (R^2= 0.16). The close agreement between observed and simulated variables for both calibration and evaluation experiments means that the model can be used to predict performance of maize across different water and nitrogen levels. The outcomes of simulations resulted across all treatments indicating that the efficiency and robustness of the model are quite adequate and the model can be used in the environments under study. The results of many studies revealed the calibration and parameterization of CERES-maize model to test irrigation management for future use in different parts of the world (Hoogenboom et al., 2004; Nouna et al., 2000; Panda et al., 2004). A comparison of the performance evaluation of the EPIC and CSM-CERES-maize models was done using maize variety trial data. The results indicated that variations between the simulations of CSM-CERES-maize and measurements were less than 3% for calibration and equal to 8% for validation (Bao et al., 2017).

By contrast, similar study in Turkey found that WUE was under-estimated by 1.5% under non water stress treatments while it was over predicted by 1.4 and 1.7% for 75% and 50% of crop water requirement, respectively (Gercek & Okant, 2010). Numerous studies have shown previously that simulation of soil water content and maize yield is accurate enough for irrigated conditions, while recent studies using CSM-CERES-Maize as part of DSSAT package have raised more concerns about the accuracy of the model in its simulation mode under deficit-irrigation conditions in semi-aired environments (DeJonge et al., 2011; Dokoohaki et al., 2017; Mubeen et al., 2013). Also, CERES-Maize calibration and validation for maize for Delhi production environment open the way for use of the model for inputs and resource management, yield forecasting and climate change impact analysis. The performance of the model after its validation was satisfactory and the results were within significant limits and were similar to the results of Esmaeilian et

al. (2014) and Ma et al. (2006). Overall, the results of this study revealed that model was able to simulate the maize grain and biomass yield accurately for full and deficit-irrigation treatments under a semi-arid condition.

CONCLUSION

This manuscript evaluated the CERES-maize model in a split plot experiment including four factors of irrigation levels and nitrogen fertilizer application. For each of the 12 combination of irrigation and fertilizer treatments, the CERES-maize model was evaluated for grain yield and biomass. The CERES-maize model was calibrated, evaluated and it estimated yield and biomass under the three N application rates with reasonable accuracy. It was observed that the CERES-maize model was more accurate in predicting the maize yield under full and 75% of FC irrigation as compared to the rainfed and 50% of FC irrigation. Finally, from the results of field data and modeling, it can be recommended that CERES-maize model can be applied to estimate the maize yield with acceptable accuracy under dynamic water and nitrogen regimes in the semi-arid environment.

REFERENCES

Abedinpour, M., Sarangi, A., Rajput, T. B. S., & Singh, M. (2014). Prediction of maize yield under future water availability scenarios using Aquacrop model. *The Journal of Agricultural Sci*ence*, 152*(4), 558–574.

Amaral, T. A., Andrade, C. L. T., Hoogenboom, G., Silva, D. F., Garcia, A., Garcia, Y., & Noce, M. A. (2015). Nitrogen management strategies for maize production systems: Experimental data and crop modeling. *International Journal of Plant Production, 9*(1), 1-24.

Bao, Y., Hoogenboom, G., Mcclendon, R., & Vellidis, G. (2017). A comparison of the performance of the CSM-CERES-maize and EPIC models using maize variety trial data. *Agricultural Systems, 150*, 109-119.

DeJonge, K. C., Ascough, J. C., Ahmadia, M., Andalesc, A. A., & Arabia, M. (2012). Global sensitivity and uncertainty analysis of a dynamic agroecosystem model under different irrigation treatments. *Ecological Modelling, 231*, 113–125.

Dogan, E., Clark, G. A., Rogers, D. H., Martin, V., & Vanderlip, R. L. (2006). On-farm scheduling studies and CERES-Maize simulation of irrigated corn. *Applied Engineering in Agriculture, 22*(4), 509–516.

Dokoohaki, H., Gheysari, M., Mousavi, S. F., & Hoogenboom, G. (2017). Effects of different irrigation regimes on soil moisture availability evaluated by CSM-CERES-Maize model under semi-arid condition. *Ecohydrology & Hydrobiology, 17*(3), 207-216.

Esmaeilian, Y., Ramroudi, M., Galavi, M., Amiri, E., & Asgharipour, M. R. (2014). Performance evaluation of CERES-Maize in simulating maize yield and WUE under water and nitrogen managements in Northern Iran. *International Journal of Biosciences, 4*(11), 10-20.

Gerçek, S., & Okant, M. (2010). Evaluation of CERES-maize simulation model results with measured data using water pillow. *African Journal of Agricultural Research, 5*(8), 606-613.

Hargreaves, J. N. G., & McCown, R. L. (1988). CERES-Maize: A Versatile Interactive Version of CERES-Maize. *CSIRO Tropical Agronomy Technical. Mem/CSIRO Division of Tropical Crops and Pastures*. St. Luica, QLD, Australia.

Hodges, T., Botner, D., Sakamoto, C., & Haug, J. H. (1987). Using the CERES-Maize model to estimate production for the US Cornbelt. *Agricultural and Forest Meteorology, 40*(4), 293–303.

Hoogenboom, G. (2004). *Genetic Coefficients CERES -maize/sorghum/millet*. In *South Asia Regional Training Workshop on "Crop Simulation Modeling"* (pp. 255-269). Multiple Cropping Center, Chiang Mai University, Thailand.

Hoogenboom, G., Porter, C. H., Shelia, V., Boote, K. J., Singh, U., White, J. W., ... & Asseng, S. (2017). *Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.7 (https://DSSAT. net)*. DSSAT Foundation, Gainesville, Florida.

Jianmei, J., Huanjie, C., Jianqiang, H., & Hongjie, W. (2014). Performance Evaluation of CERES-wheat model in Guanzhong Plain of Northwest China. *Agricultural Water Management, 144*, 1–10.

Khaliq, T., Ahmad, A., Hussain, A., & Hoogenboom, G. (2007, November). Modeling Nitrogen Use Efficiency of Maize Cropping Systems in Pakistan. In *Proceedings of the International Annual Meeting of ASA-CSSA-SSSA* (pp. 4-8). New Orleans, Louisiana.

Ló Pez-Cedró, N., Boote, K. J., Pineiro, J., & Sau, F. (2008). Improving the CERES-maize model ability to simulate water deficit impact on maize production and yield components. *Agronomy Journal, 100*(2), 296-307.

Ma, L., Hoogenboom, G., Ahuja, L. R., Ascough II, J. C., & Saseendran, S. A. (2006). Evaluation of the RZWQM-CERES-Maize hybrid model for maize production. *Agricultural System, 87*(3), 274-295.

Mubeen, M., Ahmad, A., Wajid, A., Khaliq, T., & Bakhsh, A. (2013). Evaluating CSM-CERES-Maize model for irrigation scheduling in semi-arid conditions of Punjab, Pakistan. *International Journal of Agriculture and Biol*ogy*, 15*(1), 1–10.

Nouna, B. B., Katerji, N., & Mastrorilli, M. (2000). Using the CERES-maize model in a semi-arid Mediterranean environment. New modeling of leaf area and water stress functions. *European Journal of Agronomy, 19*(2), 115-123.

Panda, R. K., Behera, S. K., & Kashyap, P. S. (2004). Effective management of irrigation water for maize under stressed conditions. *Agricultural Water Management, 66*(3), 181–203.

Wang, S. F., Li, H. L., Yang, Y. H., Wang, H. J., Yang, Y. M., & Jia, Y. G. (2012). Using DSSAT model to assess spring wheat and maize water use in the arid oasis of Northwest China. *Journal of Food Agriculture and Environment, 10*(1), 911-918.

Pertanika J. Sci. & Technol. 26 (4): 1619 - 1635 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Forecasting Share Prices Accurately For One Month Using Geometric Brownian Motion

Nur Aimi Badriah, N.[1], Siti Nazifah, Z. A.[2] and Maheran, M. J.[3]*

[1]Evyap Sabun (M) Sdn Bhd, PLO 70, Jalan Nibong 4, Tanjung Langsat Industrial Comlex, 81700 Pasir Gudang, Johor, Malaysia
[2]Universiti Teknologi MARA (UiTM) Melaka, Kampus Jasin, 77300 Merlimau, Melaka, Malaysia
[3]Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia

ABSTRACT

Investment funds are growing in Malaysia since people are more knowledgeable about investments and aware of investment opportunities in order to secure good savings for the future. These investments include unit trusts, gold, fixed deposits, stock prices and property investments. It is essential for individuals or organisations to know the value of future share prices of their investment portfolio in order to predict the profit or loss in the future. The purpose of study is to identify the best duration of historical data and forecast days in order to accurately forecast share prices. The study uses Geometric Brownian Motion model in forecasting share prices of companies in Bursa Malaysia. This study focused on 40 listed companies in Bursa Malaysia from the top gainers list. It was found that 65 historical days could forecast the share prices for 21 days accurately.

Keywords: Forecast, investment, financial mathematics, Geometric Brownian Motion, share prices

INTRODUCTION

Bursa Malaysia acts as a stock exchange in Malaysia and operates a fully integrated exchange services that offers products like equities, derivatives, and exchange-related sevices, for example clearing, trading, depository and settlement. According to Bursa Malaysia (Bursa Malaysia, 2016), there are two markets in Bursa Malaysia that are Main Market and Ace Market, which currently have 808 companies and 113 companies, respectively as of 14th December 2016.

Siti Nazifah and Maheran (2012a) stated that Bursa Malaysia was a place for people who wanted to invest in stocks. Bourses are

ARTICLE INFO

Article history:
Received: 27 November 2017
Accepted: 11 June 2018
Published: 24 October 2018

E-mail addresses:
maheran@tmsk.uitm.edu.my (Maheran, M. J)
aiminasharrudin@evyap.com.tr (Nur Aimi Badriah, N.)
sitinazifah@melaka.uitm.edu.my (Siti Nazifah, Z.A.)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

places where stocks, securities, options, and commodities are bought and sold under certain rules and regulations with organized guidelines. Some of the stocks that are traded in the Bursa Malaysia include share prices, warrants, indices, bonds and swaps. This study only focuses on the share prices.

Mohd Alias (2007) stated there were quantitative forecasting technique and qualitative forecasting technique. The quantitative forecasting technique is used based on the past available data or known as historical data. The qualitative technique depends on the accumulated experience of experts to make the possible outcome of future events (Lee, Song, & Mjelde, 2008; Omar & Halim, 2016).

Another method that is widely used to forecast share price is the Geometric Brownian Motion (GBM). Ladde and Wu (2009); Siti Nazifah and Maheran (2012a) and Wilmott (2007), used GBM to forecast share prices. Wilmott (2007) used the GBM in forecasting stock prices in the Buenos Aires Stock Exchange but did not discuss the accuracy of the model. Besides, Ladde and Wu (2009), proposed a modified version of the GBM model using different data partitioning, with and without jumps. At the end of the study, they found that data partitioning gave a good fit and the model with jumps gave a minimum mean square error (MSE) and minimum variance which implied that the best simulation produced better results than without jumps. They also stated that the nature of the GBM model did not reflect the true stock price movement except in very short terms. Meanwhile, Siti Nazifah and Maheran (2012a), used simple volatility in applying the GBM model in their study. They found the result to be highly accurate for a maximum of a two week investment period of time. In short, they had shown that the GBM model was highly accurate in forecasting the share price for short term investment.

LITERATURE REVIEW

The GBM model is a sort of continuous time model which is widely used to explain the stock price time series (Straja, 1997; Shakila, Noryati, & Maheran, 2017).

According to Wilmott (2007) and Bell (2016), the Brownian motion at time t denoted as $X(t)$ can be defined as the limiting process for this random walk as the time step goes to zero. The important properties of Brownian motion are as follows (Wilmott, 2007):

(i) Finiteness - any increment of scaling over time step in either a limit without any motion at all or a random walk leads to infinity in a finite time.

(ii) Continuity - The paths are continuous.

(iii) Markov property - The conditional distribution of $X(t)$ given information up until time $X(\tau)$ depends only on $X(\tau)$.

(iv) Martingale property - The information until time $\tau < t$, the conditional expectation of $X(t)$ is $X(\tau)$.

(v) Quadratic variation - Divide up the time 0 to t in a partition with n+1 partition points, $t_i = \frac{it}{n}$ then $\sum_{j=1}^{n} (X(t_j) - X(t_{j-1}))^2 \to t$

(vi) Normality - Normally distributed with mean zero and variance $t_i - t_{i-1}$.

Rathnayaka, Jianguo and Seneviratna (2014) prefered GBM as a model for share prices because the share prices fluctuated and showed linear and non-linear behaviours. GBM model deals with randomness, drift, volatility and returns (Siti Nazifah, 2014).

The returns at time i, where S_i is the price at the time i as a random variable drawn from a normal distribution is as follows (Siti Nazifah & Maheran, 2012a; Wilmott, 2007):

$$R_i = \frac{S_{i+1} - S_i}{S_i} \quad , \; i = 1, 2, 3, ..., n \tag{1}$$

The equation (1) can be defined as a growth of stock value at the time i. If the rate of return is negative, it shows that the closing price on the day t_{i+1} is lower than the closing price on day t_i. Conversely, a positive value indicates profit in the investment. The drift rate is the average rate of return at which the price of asset rises in a period of time (Bell, 2016; Wilmott, 2007). The formula of the drift rate μ is shown in equation (2).

$$\mu = \frac{1}{m\delta t} \sum_{i=1}^{m} R_i \tag{2}$$

where m is the number of daily returns while δt represents the timestep which is equal to the approximate number of 1/252. The value 252 is the approximate number of days trading in a year.

Based on Wilmott (2007), Abdul Rahim, Zahari and Shariff (2017), and Blyth (2014), the volatility σ refers to the fluctuation or movement of the share prices, where the prices of security move up or move down. There are four types of volatility mentioned by Wilmott (2007), but this study will use simple volatility because it is the easiest way to represent deviation from the mean value, as follows:

$$\sigma = \sqrt{\frac{1}{(m-1)\delta t} \sum_{i=1}^{m} (R_i - \overline{R})^2} \tag{3}$$

The quantity $\overline{R}$ is the mean of the returns. Siti Nazifah and Maheran (2012b) study the best volatility measurement in stock price forecasting in GBM. The result indicates higher lows-close volatility, which is closer to the actual prices proven by the MAPE value.

In this study, $S(t)$ is the stock price at time t, $X(t)$ is the random value at time t, the volatility is σ and the drift is μ will form the stochastic model as follows (Wilmott, 2007):

$$dS = \mu S dt + \sigma S dX \text{ with } F(S) = \ln S \quad (4)$$

By using Ito's Lemma (Wilmott, 2007):

$$dF = \frac{dF}{dS} dS + \frac{1}{2} \frac{d^2F}{dS^2} dS^2$$

$$dF = \frac{dF}{dS} [\mu S dt + \sigma S dX] + \frac{1}{2} \frac{d^2F}{dS^2} [\mu S dt + \sigma S dX]^2 \quad (5)$$

and

$$dX.dX = dt$$

$$dt.dt = dX.dt = dt.dX = 0$$

Simplify (5) to get

$$\frac{dF}{dS} = \frac{1}{S}, \frac{d^2F}{dS^2} = -\frac{1}{S^2} \quad (6)$$

Then by substituting $\frac{dF}{dS} = \frac{1}{S}, \frac{d^2F}{dS^2} = -\frac{1}{S^2}$ which is from (4) into (6), the following is obtained:

$$dF = \frac{1}{S} [\mu S dt + \sigma S dX] + \frac{1}{2} \sigma^2 S^2 (-\frac{1}{S^2}) dt$$

$$dF = \frac{1}{S} [\mu S dt + \sigma S dX] - \frac{1}{2} \sigma^2 dt$$

$$dF = \frac{1}{S} \mu S dt + \frac{1}{S} \sigma S dX - \frac{1}{2} \sigma^2 dt$$

$$dF = \mu dt + \sigma dX - \frac{1}{2} \sigma^2 dt$$

$$dF = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dX.$$

Integrate both sides,

$$\int dF = \int \left(\mu - \frac{1}{2}\sigma^2 \right) dt + \int \sigma dX$$

$$\ln S = \left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma (X(t) - X(0)) + c$$

$$e^{\ln S} = e^{\left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma (x(t) - x(0)) + c}$$

$$S = e^{\left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma (x(t) - x(0))} . e^{c}$$

Here $e^c = S(0)$ and with that, the equation of forecasting share prices is produced,

$$S(t) = S(0) e^{(\mu - \frac{1}{2}\sigma^2) t + \sigma (X(t) - X(0))} \quad (7)$$

where $S(t)$ is the price of stock at time *t*. In this study, this equation was used to forecast share prices of companies in Bursa Malaysia.

To know the accuracy of the forecast value, an accuracy test is essential. MAPE is widely used in order to evaluate the performance of the algorithm. Basically, MAPE is the average absolute error between the actual values and predicted values in percentage. Some of the researchers who used MAPE were Aslina and Maheran (2011), Gupta and Dhingra (2012), Hadavandi et al. (2010), Siti Nazifah and Maheran (2012a, 2012b), and Sonono and Mashele (2015).

In this study, the accuracy of the actual and forecast values can be measured by using MAPE. The formula for the MAPE, *E* is shown in equation (2.12) (Reid, 2013; Siti Nazifah & Maheran, 2012a),

$$E = \frac{\sum \left| \frac{Y_t - F_t}{Y_t} \right|}{n} \times 100\% \quad (8)$$

which involves time period *t*, the number of period forecast *n*, actual value in the time period at time *t*, Y_t and forecast value in time period *t*, F_t. In order to identify the accuracy of the model, Lawrance, Kimberg and Lawrence (2009) proposed a scale of judgement based on MAPE as in Table 1. The smaller the MAPE value, the more accurate is the forecasting model.

Table 1
A forecast accuracy scale of judgement

MAPE	Forecast Accuracy
Less than 10%	Highly accurate
11% to 20%	Good forecast
21% to 50%	Reasonable forecast
51% and above	Inaccurate forecast

Source: (Lawrence, Kimberg, & Lawrence, 2009)

METHODS

Data Collection

In Bursa Malaysia, the share prices are traded daily in two sessions on Monday to Friday. The first session is from 8.30 am until 12.30 pm while the second session is at 2.00 pm until 5.00 pm. The share prices change every time. To forecast the share price in one month's time, the data of the closing price were taken from the Bursa Malaysia list.

There are many companies that are listed on Bursa Malaysia. Based on Kanniainen and Keuschnigg (2003), a minimum number of companies to set up a portfolio was 5. Thus, the share prices of 5 companies were selected for each sector which meant there were 40 companies in eight sectors of Bursa Malaysia. The chosen companies were picked based on the Top Gainers list on the Bursa Malaysia website as of 13th November 2016.

In this study, the data were taken from 1st July 2015 until 30th September 2016. For forecasting the share prices, the historical data from 1st July 2015 until 30th June 2016 for each selected 40 companies were considered as input data. Data from 1st July 2016 until 30th September 2016 were used to make a comparison with the forecast value. The historical data were important as they determined the accuracy of the value and it was useful to see the pattern of future stock prices. As for the missing values, the average of the share price of the day before and the day after was calculated.

From the historical data, the study analysed its characteristics and performance of share prices in every company. The characters of the historical data were critical in determining the forecast value and also the pattern for future share prices.

The data played an important role in forecasting as accurately as possible when applying the GBM model. Thus, the data had to be sufficient and relevant in order to fulfill the objectives. According to the Wilmott (2007), the share price followed a random walk since the data changed every day and in fact every second.

Forecasting and Calculating Error

First and foremost, the value of return was calculated where the equation of rate of return (1) R_i for day 1 was used and next, the average of return distribution $\overline{R}$ was obtained. The

value m = 65 was the number of returns in the sample. Nur Aimi (2017) had tested with different number of return m=262, 130, 88 and 65. It is found that the error is small when m=65. The value of R_i and $\delta t = \frac{1}{252}$ for United Plantation Berhad is shown in Table 2 below.

Table 2
Return and average return value of for United Plantation Berhad with 65 input data

Number of Data	Date	Share Price, S_i	Return, R_i	$(R_i - \overline{R})^2$
	31/03/2016	26.50	-0.00528	0.0000295194852406
1	01/04/2016	26.36	-0.00228	0.0000058871300060
2	04/04/2016	26.30	0	0.0000000225494802
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
64	29/06/2016	26.62	0.00	0.0000000225494802
65	30/06/2016	26.62	0.00	0.0000000225494802
	Total Return		0.00976	-
	Average Return		0.00015	-
	Total $(R_i - \overline{R})^2$		-	0.010477362

The value of $\delta t = \frac{1}{252}$ that 252 was the total trading days of share prices in a year was set (Wilmott, 2007). The expected return or growth rate of an asset is called drift rate μ and was obtained by using equation (2). After finding the drift, calculate the volatility, σ using the equation (3).

The values of drift and volatility were obtained for all the 40 companies and are shown in Table 5. In order to find the forecast value of share prices, the equation (7) was used.

The accuracy of the forecast value was measured by using MAPE. Before getting the value of MAPE, the absolute percentage error, ε was calculated as shown below:

$$\varepsilon_1 = \left| \frac{Actual - Forecast}{Actual} \right| \times 100\%$$

The summation of 21 absolute percentage error of 21 forecast days was then divided by 21 to get the value of MAPE, *E* as in equation (8). The study crried out the same forecast calculation for all 40 companies.

The Microsoft Excel

The calculation in the study was done by using the Microsoft Excel. Figure 1 shows an example of the spreadsheet for the United Plantation Berhad company. The forecast value changed every time the enter key was used or the file was saved due to the presence of the generated random variable in the forecast formula by Microsoft Excel.

	G	H	I	J	K	L	M	N	O	P
1				t	No of Data	Date	Actual	Forecast	Absolute Percentage Error	MAPE
2	m	65				30/06/2016	26.62			
3	m-1	64		0.003968254	1	01/07/2016	26.62	26.95482173	1.257782615	5.817564111
4	Rbar	0.000150165		0.007936508	2	04/07/2016	26.62	28.69971381	7.812598835	
5	dt	0.003968254		0.011904762	3	05/07/2016	26.62	28.76311195	8.050758649	
6	Drift,μ	0.037841541		0.015873016	4	06/07/2016	26.62	27.84411906	4.598493858	
7	Volatility,σ	0.203112314		0.01984127	5	07/07/2016	26.62	25.60801372	3.801601355	
8				0.023809524	6	08/07/2016	27	25.49110179	5.588511894	
9	x(0)	0.409354393		0.027777778	7	11/07/2016	27	27.62721208	2.323007715	
10				0.031746032	8	12/07/2016	26.8	26.96402042	0.612016478	
11				0.035714286	9	13/07/2016	26.8	25.14158973	6.188098019	
12				0.03968254	10	14/07/2016	26.54	28.72445591	8.230805994	
13				0.043650794	11	15/07/2016	26.44	29.7637083	12.57075757	
14				0.047619048	12	18/07/2016	26.98	24.74141828	8.297189492	
15				0.051587302	13	19/07/2016	26.98	29.86550533	10.69497897	
16				0.055555556	14	20/07/2016	27.1	27.68007847	2.140510973	
17				0.05952381	15	21/07/2016	27.1	25.77675963	4.882805803	
18				0.063492063	16	22/07/2016	27.1	25.83793572	4.657063755	
19				0.067460317	17	25/07/2016	26.78	28.23718139	5.441304658	
20				0.071428571	18	26/07/2016	26.78	29.52669486	10.25651555	
21				0.075396825	19	27/07/2016	27.06	27.450646	1.443628966	
22				0.079365079	20	28/07/2016	27.06	28.36727699	4.831031003	
23				0.083333333	21	29/07/2016	26.54	28.79308256	8.489384186	
24									122.1688463	

Figure 1. Calculation of Forecast Share Price for United Plantation Berhad by using Microsoft Excel

Figure 1 shows how to evaluate the forecast share price by using an Excel template. Cell N3 is formulated to forecast the value of the share price as follows:

M2*EXP((H6-(1/2)*(H7^2))*J3 + (H7*(RAND()-H9)))

where the value of M2 was 26.62 meanwhile the forecast value at cell N3 was 26.9548. Random number was generated by Microsoft Excel using *RAND()* command.

The Accuracy for Duration of Data and Forecast Days

The aim at this stage was to find the best duration of data used in the GBM model for better forecasting. The difference in length of duration of input data and a number of forecast days used was essential in getting the smallest possible value of MAPE. The smaller the

MAPE, the more accurate it was to the actual price. Table 3 shows the starting and ending dates of daily data used for the duration for input data.

Table 3
The number of data for input duration and dates of daily data

Starting Date	Ending Date	Duration	Number of Data
1st July 2015	30th June 2016	12 months	262
1st January 2016	30th June 2016	6 months	130
1st March 2016	30th June 2016	4 months	88
1st April 2016	30th June 2016	3 months	65

The study was conducted on these durations of input data because the drift and volatility value were different at different duration. This certainly affected the forecast value.

The values of drift and volatility for a different number of historical data of 262 days, 130 days, 88 days and 65 days were calculated and presented in Table 5. With these different value of drift and volatility, the share price was forecast for all the 40 companies and for all the duration of historical data.

The MAPE was obtained in studying the error. The values MAPE are presented in Table 6 for the 4 different number of historical data which were 262 days, 130 days, 88 days, and 65 days for all the 40 companies. The three different numbers of forecast days, which were 66 days, 44 days, and 21 days. 66, 44 and 21 days represented 3 months, 2 months, and 1 month of forecast values respectively.

Table 4 shows the starting and ending dates of forecast duration. It also shows the duration and the number of forecast days.

Table 4
The number of data for forecast duration and dates of daily data

Starting Date	Ending Date	Duration	Number of Data
1st July 2016	30th September 2016	3 months	66
1st July 2016	31st August 2016	2 months	44
1st July 2016	29th July 2016	1 months	21

Knowing the forecast price of the shares for more than 1 month can definitely help the financial advisors and decision makers to plan and to invest better. The study obtained the duration of days that had the highly accurate forecast values when using equation (7). The results of this experiment can be found in Table 7.

RESULT AND DISCUSSION

This section discusses the findings of the study. This study experimented the duration of time that could give the smallest possible value of MAPE. The different durations of time were 262, 130, 88, and 65 days as in Table 4 for input data, while the different intervals of time for forecast data were 66, 44, and 21 days as presented in Table 5. The difference in the amount of historical data resulted in the difference in the value of drift and volatility. Hence, Table 5 below shows the value of the drift, μ and volatility, σ according to the respective number of historical data.

Table 5
The values of drift and volatility for different historical days

Sector	Companies Name	Drift and volatility for different no. of historical data							
		262		130		88		65	
		$\boldsymbol{\mu}$	σ	μ	σ	μ	σ	μ	σ
Plantation	United Plantation Berhad	0.01	0.19	0.11	0.17	0.08	0.18	0.4	0.20
	Kuala Lumpur Kepong Berhad	0.08	0.16	0.04	0.13	-0.07	0.10	-0.2	0.10
	Chin Teck Plantation Berhad	-0.13	0.18	-0.11	0.17	-0.09	0.18	-0.23	0.20
	Kim Loong Resources Berhad	0.18	0.23	0.18	0.23	0.21	0.24	-0.33	0.20
	Kwantas Corp. Berhad	-0.25	0.21	-0.25	0.21	-0.25	0.21	-0.37	0.20
Consumer Product	Nestle (Malaysia) Berhad	0.06	0.10	0.08	0.08	0.09	0.10	0.08	0.06
	Latitud Tree Holdings Berhad	-0.06	0.31	-0.60	0.31	-0.36	0.30	-0.04	0.25
	Kawan Food Berhad	0.43	0.32	0.07	0.29	0.11	0.26	0.31	0.26
	Lii Hen Industries Berhad	0.51	0.67	0.36	0.40	0.69	0.42	1.65	0.41
	Poh Huat Resources Holdings Berhad	-0.06	0.61	-0.50	0.32	-0.15	0.30	0.27	0.28
Industrial Product	Top Glove Corporation Berhad	-0.10	0.62	-1.67	0.77	-0.50	0.29	-0.28	0.24
	Kossan Rubber Industries Berhad	0.09	0.32	-0.54	0.33	0.18	0.24	0.55	0.19
	TA Ann Holdings Berhad	-0.10	0.34	-0.72	0.43	-1.25	0.46	-1.49	0.47
	Lysaght Galvanized Steel Berhad	-0.11	0.26	-0.15	0.20	-0.07	0.18	-0.05	0.20
	Focus Lumber Berhad	0.40	0.45	-0.75	0.45	-0.74	0.42	-0.32	0.43
Trading and Services	Pos Malaysia Berhad	-0.33	0.45	0.16	0.55	1.03	0.61	0.23	0.35
	Luxchem Corporation Berhad	0.41	0.35	-0.21	0.28	-0.65	0.22	-0.63	0.22
	Ipmuda Berhad	-0.19	0.46	-0.16	0.27	-0.14	0.12	-0.15	0.10
	Harrisons Holdings (Malaysia) Berhad	-0.17	0.20	0.02	0.16	-0.02	0.16	-0.01	0.16
	Petronas Dagangan Berhad	0.14	0.16	-0.11	0.16	-0.18	0.10	-0.11	0.10
Properties	IGB Corporation Berhad	-0.14	0.16	0.14	0.17	0.08	0.18	-0.06	0.18
	Selangor Dredging Berhad	0.01	0.30	0.03	0.25	0.05	0.19	0.02	0.22
	Crescendo Corporation Berhad	-0.43	0.25	-0.37	0.23	-0.43	0.24	-0.60	0.18
	Mah Sing Group Berhad	-0.06	0.25	0.05	0.23	0.31	0.21	0.10	0.21
	Amcorp Properties Berhad	0.06	0.39	-0.07	0.19	-0.06	0.15	-0.17	0.15

Table 5 *(continue)*

Sector	Companies Name	Drift and volatility for different no. of historical data							
		262		130		88		65	
		μ	σ	μ	σ	μ	σ	μ	σ
Construction	WCT Holdings Berhad	0.13	0.38	-0.09	0.29	-0.12	0.30	-0.50	0.32
	Kimlun Corporation Berhad	0.35	0.32	0.58	0.32	0.52	0.28	0.05	0.25
	Pesona Metro Holdings Berhad	-0.43	0.36	-0.23	0.24	-0.10	0.20	-0.30	0.19
	Zelan Berhad	-0.66	0.41	-0.88	0.38	-0.95	0.39	-1.31	0.41
	Ahmad Zaki Resources Berhad	0.03	0.36	0.09	0.30	0.36	0.33	-0.27	0.29
Finance	Aeon Credit Services M Berhad	-0.03	0.19	0.18	0.18	0.25	0.18	0.05	0.14
	Affin Holdings Berhad	-0.22	0.20	-0.18	0.16	-0.03	0.14	-0.36	0.13
	ELK-Desa Resources Berhad	-0.09	0.22	-0.06	0.12	-0.09	0.13	-0.12	0.11
	LPI Capital Berhad	0.13	0.11	-0.02	0.10	0.19	0.08	0.11	0.09
	Syarikat Takaful Malaysia Berhad	0.05	0.24	0.10	0.22	0.20	0.24	-0.03	0.20
Technology	Globetronics Technology Berhad	-0.49	0.38	-1.18	0.47	-1.35	0.51	-1.74	0.57
	JCY International Berhad	-0.10	0.36	-0.64	0.32	-0.36	0.28	-0.70	0.24
	Grand-FLO Berhad	-0.11	0.32	-0.11	0.30	-0.01	0.33	0.06	0.34
	Notion VTEC Berhad	0.07	0.42	-0.18	0.29	-0.09	0.23	-0.17	0.23
	Key Asic Berhad	1.63	1.58	0.44	0.88	-0.21	0.73	-0.99	0.73

Table 5 indicates the value of drift and volatility for different duration of input data which are 262 days, 130 days, 88 days and 65 days. Before setting which drift and volatility that could be used for the next step, the value of MAPE for different forecast days were taken into consideration. The different interval of forecast share price was calculated with the different interval of input data to get the MAPE results as in Table 6.

Table 6 presents the value of the MAPE of forecast prices when compared to the actual prices when using the GBM model for 40 companies. This means that the randomness of share prices are being considered in the equation used. Based on the average value obtained, the data with 65 historical days and 21 forecast days were the best data to be taken since the average MAPE was the lowest (8.74%) compared to the rest. Table 7 presents the actual and forecast values on the 21st day, which was on 29th July 2016 with MAPE values for the whole one month of forecasting.

In Table 7, the value of forecast and actual price for all the 40 companies did not show much difference. For example, the actual value of share price of the United Plantation Berhad company on 29th July 2016 was RM26.54, the forecast value was RM28.7931 which had a difference of RM2.2531. With MAPE, the value was 5.82% of the accuracy of GBM

Table 6
Values of MAPE for all companies involved with respective number of historical data and number of forecast day

Sector	Companies Name	MAPE Value of different number of Historical Data											
		262			130			88			65		
		Number of Forecast Days											
		66	44	21	66	44	21	66	44	21	66	44	21
Plantation	United Plantation Berhad	5.65	6.79	7.91	6.18	5.33	4.40	5.43	9.83	5.86	5.41	9.54	5.82
	Kuala Lumpur Kepong Berhad	4.29	4.23	6.69	4.38	3.41	4.70	7.02	2.97	3.54	6.50	3.47	3.29
	Chin Teck Plantation Berhad	11.12	6.98	4.27	9.04	7.27	7.46	7.39	5.68	7.48	10.07	4.68	6.01
	Kim Loong Resources Berhad	5.82	4.96	5.83	6.07	7.21	5.37	5.91	6.46	6.91	8.16	4.15	8.43
	Kwantas Corp. Berhad	7.25	6.35	8.55	6.37	6.08	5.34	7.41	7.20	6.83	5.92	6.90	10.95
Consumer Product	Nestle (Malaysia) Berhad	6.24	2.67	5.37	3.02	4.48	2.39	2.56	4.23	4.62	4.04	3.31	2.32
	Latitud Tree Holdings Berhad	11.07	8.61	7.88	8.06	8.37	13.58	7.68	9.25	16.01	15.06	9.81	6.67
	Kawan Food Berhad	8.36	9.49	15.24	10.45	8.75	14.58	11.63	6.81	6.44	13.77	10.79	7.51
	Lii Hen Industries Berhad	30.16	43.64	22.44	9.93	12.14	10.64	27.24	9.44	12.07	21.55	11.41	9.82
	Poh Huat Resources Holdings Berhad	17.52	16.16	27.75	11.18	10.49	8.27	8.63	9.36	10.51	14.07	13.56	6.54
Industrial Product	Top Glove Corporation Berhad	19.70	36.29	15.74	24.63	33.62	29.93	13.01	7.91	9.51	8.45	5.96	16.92
	Kossan Rubber Industries Berhad	11.54	13.36	9.89	9.62	14.96	12.31	6.61	6.58	8.92	8.75	9.69	7.99
	TA Ann Holdings Berhad	10.21	11.61	9.40	27.95	24.28	17.14	33.15	17.55	15.04	15.96	14.23	12.60
	Lysaght Galvanized Steel Berhad	10.46	10.75	10.31	6.69	10.25	6.51	5.06	5.59	7.58	5.38	5.24	6.02
	Focus Lumber Berhad	16.92	14.20	13.53	12.22	23.75	19.85	14.45	15.84	15.13	17.17	16.72	14.62
Trading and Services	Pos Malaysia Berhad	16.21	12.65	20.51	16.64	26.56	14.95	25.53	18.08	14.83	13.64	8.77	9.50
	Luxchem Corporation Berhad	18.65	11.82	15.27	12.77	15.06	8.06	20.13	20.29	8.04	13.47	20.31	11.95
	Ipmuda Berhad	17.31	27.28	19.32	13.54	9.13	11.83	3.24	5.80	4.36	3.95	5.86	2.64
	Harrisons Holdings (Malaysia) Berhad	8.08	14.44	5.02	6.23	4.57	4.12	5.52	4.98	4.34	9.99	4.98	5.29
	Petronas Dagangan Berhad	5.77	4.82	6.32	5.84	5.58	3.64	2.88	3.78	3.02	3.68	2.65	2.23

Table 6 *(continue)*

Sector	Companies Name	MAPE Value of different number of Historical Data											
		262			130			88			65		
		Number of Forecast Days											
		66	44	21	66	44	21	66	44	21	66	44	21
Properties	IGB Corporation Berhad	4.29	12.48	6.05	4.32	7.81	8.16	8.04	4.89	4.73	11.53	7.58	5.15
	Selangor Dredging Berhad	7.59	7.26	11.04	6.59	9.19	5.33	7.41	7.11	5.95	5.52	15.00	5.43
	Crescendo Corporation Berhad	15.86	11.01	10.36	6.72	7.43	6.36	7.20	7.12	7.73	11.89	5.48	6.38
	Mah Sing Group Berhad	14.61	12.71	7.60	11.19	12.19	8.25	5.87	6.79	7.53	7.59	6.44	7.22
	Amcorp Properties Berhad	13.59	12.54	10.75	9.34	8.31	5.15	7.12	8.35	4.93	6.65	8.07	6.10
Construction	WCT Holdings Berhad	13.17	9.62	6.76	8.72	7.88	6.31	10.90	10.72	6.74	18.70	13.10	9.38
	Kimlun Corporation Berhad	10.38	8.73	8.91	13.28	10.78	10.93	11.40	8.25	12.12	9.80	8.33	10.96
	Pesona Metro Holdings Berhad	12.59	17.17	9.44	11.78	9.98	8.03	12.87	15.69	7.86	12.06	12.88	6.04
	Zelan Berhad	27.76	16.53	9.69	23.09	17.10	25.63	37.41	17.58	17.51	36.29	29.11	24.06
	Ahmad Zaki Resources Berhad	14.79	13.02	14.47	9.54	12.62	14.79	11.36	10.20	10.21	12.56	7.87	17.16
Finance	Aeon Credit Services M Berhad	11.04	12.20	5.29	9.24	8.07	5.50	4.78	4.29	10.82	8.47	10.66	7.87
	Affin Holdings Berhad	6.02	6.02	5.35	5.04	7.04	4.47	5.35	3.67	8.22	6.92	3.98	6.67
	ELK-Desa Resources Berhad	6.96	7.74	9.02	7.12	3.06	3.12	5.40	4.55	3.24	6.66	4.53	3.53
	LPI Capital Berhad	3.13	3.73	2.76	4.05	2.87	2.65	4.42	2.15	2.79	2.87	3.08	2.69
	Syarikat Takaful Malaysia Berhad	12.57	13.65	9.61	7.62	4.62	7.04	8.46	6.78	7.60	11.37	6.44	6.07
Technology	Globetronics Technology Berhad	12.63	17.84	7.15	20.01	15.27	17.01	16.71	13.06	22.32	30.90	10.10	23.14
	JCY International Berhad	17.82	10.50	11.17	8.11	14.40	9.23	17.83	9.64	6.35	7.23	5.20	6.05
	Grand-FLO Berhad	29.90	16.03	9.81	14.66	10.44	7.47	30.79	19.80	7.27	16.78	26.15	12.58
	Notion VTEC Berhad	17.92	19.81	15.13	8.37	11.16	13.48	9.61	9.22	15.12	10.32	6.85	7.42
	Key Asic Berhad	42.80	34.30	78.77	34.17	25.94	52.56	22.31	37.94	53.94	25.96	16.08	18.43
	Average MAPE	13.44	13.25	12.16	10.84	11.19	10.66	11.64	9.64	9.85	11.63	9.47	8.74

Table 7
The values of actual and forecast for day 21st with MAPE value for 1 month

Sector	Companies Name	Actual (RM)	Forecast (RM)	MAPE (%)
Plantation	United Plantation Berhad	26.54	28.7931	5.82
	Kuala Lumpur Kepong Berhad	23.12	22.7404	3.29
	Chin Teck Plantation Berhad	7.41	6.3537	6.01
	Kim Loong Resources Berhad	3.3	3.3312	8.43
	Kwantas Corp. Berhad	1.21	1.1937	10.95
Consumer Product	Nestle (Malaysia) Berhad	79.00	78.8727	2.32
	Latitud Tree Holdings Berhad	5.25	5.5730	6.67
	Kawan Food Berhad	3.62	3.7274	7.51
	Lii Hen Industries Berhad	3.05	3.6817	9.82
	Poh Huat Resources Holdings Berhad	1.49	1.6842	6.54
Industrial Product	Top Glove Corporation Berhad	4.29	5.3456	16.92
	Kossan Rubber Industries Berhad	6.66	6.9338	7.99
	TA Ann Holdings Berhad	3.37	2.4256	12.60
	Lysaght Galvanized Steel Berhad	3.43	3.1194	6.02
	Focus Lumber Berhad	1.7	1.6907	14.62
Trading and Services	Pos Malaysia Berhad	2.85	2.9343	9.50
	Luxchem Corporation Berhad	1.52	1.1872	11.95
	Ipmuda Berhad	0.90	0.9302	2.64
	Harrisons Holdings (Malaysia) Berhad	2.95	2.9563	5.29
	Petronas Dagangan Berhad	23.26	23.6916	2.23
Properties	IGB Corporation Berhad	2.55	2.6684	5.15
	Selangor Dredging Berhad	0.895	0.8186	5.43
	Crescendo Corporation Berhad	1.57	1.5147	6.38
	Mah Sing Group Berhad	1.62	1.7002	7.22
	Amcorp Properties Berhad	0.91	0.8240	6.10
Construction	WCT Holdings Berhad	1.56	1.80	9.38
	Kimlun Corporation Berhad	1.78	1.703	10.96
	Pesona Metro Holdings Berhad	0.37	0.398	6.04
	Zelan Berhad	0.165	0.093	24.06
	Ahmad Zaki Resources Berhad	0.640	0.761	17.16
Finance	Aeon Credit Services M Berhad	14.20	12.3241	7.87
	Affin Holdings Berhad	2.13	1.91	6.67
	ELK-Desa Resources Berhad	1.21	1.1838	3.53
	LPI Capital Berhad	16.06	16.3010	2.69
	Syarikat Takaful Malaysia Berhad	4.01	3.5672	6.07
Technology	Globetronics Technology Berhad	2.91	3.2668	23.14
	JCY International Berhad	0.595	0.5533	6.05
	Grand-FLO Berhad	0.22	0.2404	12.58
	Notion VTEC Berhad	0.415	0.3848	7.42
	Key Asic Berhad	0.125	0.0989	18.43
Average MAPE				8.74

by using 65 historical days and 21 days of the forecast was highly accurate. There were some companies that produced higher MAPE such as Top Glove Corporation Berhad and Globetronics Technology Berhad with MAPE of 16.92% and 23.14% respectively. However, according to the scale of judgement of forecast accuracy, both companies were still considered as good forecast. Based on Table 7, the lowest MAPE value was from Nestle (M) Berhad with 2.32% was highly accurate. This meant that the forecast value was 2.32% different from the actual value. Meanwhile, the highest MAPE value went to Zelan Berhad with 24.06% which could be considered as a reasonable forecast. Most of forecast values were highly accurate. The random variable that are being considered in equation is able to give accurate forecast share prices.

CONCLUSION

To forecast share prices, the measurement of accuracy of the forecast share price is important. An experiment was done to find the best time interval of historical data and duration of forecast days in order to get the lowest MAPE which meant the high accuracy of the forecast value. The results of this experiment showed that the duration of 65 days of historical data could forecast the share prices for 21 days or one month accurately. With the average MAPE for all companies was 8.74%, these forecast value were highly accurate. With the high accuracy of forecast share prices, the financial institutions, organizations or individual investors will be able to make better decisions.

ACKNOWLEDGEMENT

This study is funded by the Fundamental Research Grant Scheme (FRGS), Ministry of Higher Education Malaysia that is managed by the Research Management Centre (RMC), IRMI, Universiti Teknologi MARA, 600-IRMI/FRGS 5/3 (83/2016). The authors would like to thank Mr. Shaiful Bakhtiar Rodzman in managing the paper for publication.

REFERENCES

Bell, S. (2016). *Quantitative Finance for Dummies*. Chichester: John Wiley & Sons Ltd.

Blyth, S. (2014). *An Introduction to Quantitative Finance*. Oxford: OXFORD University Press.

Bursa Malaysia. (2016). Retrieved November 13, 2016, from www.bursamalaysia.com

Chai, T., & Draxler, R. R. (2014). Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250.

Gupta, A., & Dhingra, B. (2012, March). Stock market prediction using hidden markov models. In *2012 Students Conference on Engineering and Systems (SCES)* (pp. 1-4). Allahabad, Uttar Pradesh.

Hadavandi, E., Shavandi, H., & Ghanbari, A. (2010). Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. *Knowledge-Based Systems*, *23*(8), 800–808.

Kanniainen, V., & Keuschnigg, C. (2003). The optimal portfolio of start-up firms in venture capital finance. *Journal of Corporate Finance*, *9*(5), 521–534.

Ladde, G. S., & Wu, L. (2009). Development of modified Geometric Brownian Motion models by using stock price data and basic statistics. *Nonlinear Analysis, Theory, Methods and Applications*, *71*(12), e1203–e1208.

Lawrence, K. D., Klimberg, R. K., & Lawrence, S. M. (2009). *Fundamentals of Forecasting using Excel*. New York: Industrial Press.

Lazim, M. A. (2007). *Introductory Business Forecasting: A Practical Approach*. Shah Alam: Pusat Penerbitan Universiti, Universiti Teknologi MARA.

Lee, C. K., Song, H. J., & Mjelde, J. W. (2008). The forecasting of international expo tourism using quantitative and qualitative techniques. *Tourism Management*, *29*(6), 1084–1098.

Nur Aimi Badriah, N. (2017). Predicting the risk value of share prices portfolio by integrating value at risk method with Geometric Brownian Motion, (Master dissertation), Universiti Teknologi MARA, Malaysia.

Omar, A., & Jaffar, M. M. (2011, September). Comparative analysis of Geometric Brownian motion model in forecasting FBMHS and FBMKLCI index in Bursa Malaysia. In *2011 IEEE Symposium on Business, Engineering and Industrial Applications (ISBEIA)* (pp. 157-161). Langkawi, Malaysia.

Omar, N. A. B., & Halim, F. A. (2015, May). Modelling volatility of Malaysian stock market using garch models. In *International Symposium on Mathematical Sciences and Computing Research (iSMSC)* (pp. 447-452). Tapah, Malaysia.

Rahim, M. A., Zahari, S. M., & Shariff, S. S. R. (2017). Performance of Variance Targeting Estimator (VTE) under misspecified error distribution assumption. *Pertanika Journal of Science & Technology, 25*(2), 607-618.

Rathnayaka, R. M. K. T., Jianguo, W., & Seneviratna, D. M. K. N. (2014). Geometric Brownian Motion with Ito's lemma approach to evaluate market fluctuations: A case study on Colombo Stock Exchange. In *Proceedings of 2014 IEEE International Conference on Behavioral, Economic, Socio-Cultural Computing, (BESC)* (pp. 1-6). Shanghai, China.

Reid, H. M. (2013). *Introduction to Statistics*. Thousand Oaks: SAGE Publications.

Shakila, S., Noryati, A., & Maheran, M. J. (2017). Assessing stock market volatility for different sectors in Malaysia. *Journal of Science and Technology, 25*(2), 631-648

Siti Nazifah, Z. A., & Maheran, M. J. (2012a). A Review on Geometric Brownian Motion in Forecasting the Share Prices in Bursa Malaysia. *World Applied Sciences Journal*, *17*, 87–93.

Abidin, S. N. Z., & Jaafar, M. M. (2012, June). Surveying the best volatility measurements in stock market forecasting techniques involving small size companies in Bursa Malaysia. In *2012 IEEE Symposium on Humanities, Science and Engineering Research (SHUSER)* (pp. 975-979). Kuala Lumpur, Malaysia.

Siti Nazifah, Z. A. (2014). *Intergrating cluster analysis, geometric brownian motion and analytic hierarchy process in capital allocation,* (Master dissertation), Universiti Teknologi MARA, Malaysia.

Sonono, M. E., & Mashele, H. P. (2015). Prediction of stock price movement using continuous time models. *Journal of Mathematical Finance, 5*(02), 178–191.

Straja, S. R. (1997). *Stochastic modeling of stock prices*. Camden, NJ: Montgomery Investment Technology.

Wilmott, P. (2007). *Introduces quantitative finance*. Chichester: John Wiley.

Pertanika J. Sci. & Technol. 26 (4): 1637 - 1658 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Fuzzy Production Inventory Model with Stock Dependent Demand Using Genetic Algorithm (GA) Under Inflationary Environment

Pallavi Agarwal[1]*, Neeraj Kumar[2] and Ajay Sharma[1]

[1]Department of CSE, SRM University Delhi – NCR, Sonepat, Haryana 131029, India
[2]Department of Mathematics, SRM University Delhi – NCR, Sonepat, Haryana 131029, India

ABSTRACT

In this article, we want to solve the complication of production of the product for the newly launched product and integrate it with the value of time, product, and inflation value. To address such problems, we have used a linear demand rate, which is directly proportional to the stock level i.e. if the stock level is maximum, the demand will automatically increase and the inventory level will also increase. If the level of the stock decreases then the demand and inventory level will also decrease. Moreover, the production will be stopped, when the level of stock will reach level S and there is no effect of demand. The S_0 stock level is definite. The prospective of this research is to increase the total profit of the model. For which the use of the Centroid method of defuzzification is used to defuzzify the total profit. This model will be explained with the help of numerical examples and sensitivity analysis and Java and MATLAB R2015a are used to get the optimal values of this model.

Keywords: Centroid Method, EPQ model, genetic algorithm, stock-dependent demand, Supply Chain Management

INTRODUCTION

India is a country of diversity where within 50 km people is speaking different languages and they are having different requirements during time to time and it changes day by day. This kind of situation is called as the demand-reliant on the stock. In this field, so many mathematical models have already been discussed in the prevailing literature. One of these models was by Chang et al. (2004), Datta et al. (1998), Mishra (1979),

ARTICLE INFO

Article history:
Received: 09 September 2017
Accepted: 25 April 2018
Published: 24 October 2018

E-mail addresses:
pallavirupali@gmail.com (Pallavi Agarwal)
nkneerajkapil@gmail.com (Neeraj Kumar)
ajaypulast@rediffmail.com (Ajay Sharma)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

Singh et al. (2007), Singh et al. (2010), Teng and Chang (1995), and Yao et al. (2000). These conventional inventory models are installed on the root where all the obstacles are fixed in the total cost and they have some parameters or value. But the time and place are different for the place of market setup and the value of the parameters or constraints may increase or decrease according to the environment or the needs of the people. To handle these types of situations, we will assume that all the criteria are fuzzy and the resulting function will also be fuzzy.

To get an impression about this research, we refer to Chandra and Bahner (1985), Maity and Maiti (2007), Teng and Chang (2005), Singh and Singh (2010), and Wu et al. (2006). Classical inventory models neglect the effects of inflation because inflation does not affect inventory rules by any significant degree but we cannot ignore inflation because it is directly or indirectly proportional to return on investment.

Inflation should be especially measured for long period investment and forecast such as ice and cold storage industry and weather forecasts, which specifically help in the production industry (Agarwal et al., 2017). Discussing in which to deal with the actual situation of the ice and cold storage (Agra) and to find the solution to the problem. To get the solution of these problems they used GA.

To get an impression on this research, we refer to Balkhi and Benkherouf (2004), Buzacott (1975), Lieo et al. (2000), Mehta and Shah (2003), Yao and Lee (1999), and Singh and Singh (2010) who discussed GA's experiments in the complex decision-making. In addition, Kumar et al. (2016) and Sarkar et al. (1997), discussed about improving the quality of GA for single/multi-purpose continuous/discrete optimization problems during the last days. They also discussed a model with constant demand which is based on the stock level under inflation environment and GA was the solution of this kind of problems. Michalewich (1992) extends a GA, which is called Contractive Mapping Genetic Algorithm (CMGA).

In this model, they discussed that the new population would be selected when it was proven that the new population or generation was fitter compared to the old population and it had been obtained by the asymptotic convergence of the algorithm by the Banach Fixed Point Theorem.

Bessaou and Siarry (2001) discussed a GA where they were creating more than one solution for the population of solutions in the beginning, then genetic operations were done on every population. Last and Eval (2005) discussed about a model of a GA, which was based on different chrome shapes and different ages (such as child age, young age, middle age and old age) chrome. Pezzella Morgantia and Ciaschettib (2008) extended a GA that assimilated new ones with different sets to re-develop and also satisfied the job-shop scheduling problem.

Comparison is presented in the Table 1.

Table 1
Comparison is presented in the following table of the literature

S. No	Author's	EPQ Model	Inflation	Centroid method	Stock Dependent Demand	GA
1	Agarwal (2017)					Yes
2	Balkhi (2004)				Yes	
3	Bessaou and Siarry (2001).					Yes
4	Buzacott (1975)		Yes			
5	Chandra and Bahner (1985).		Yes			
6	Chang, Yao and Quyang (2004).			Yes		
7	Datta, Paul and Pal (1998).				Yes	
8	Kumar, Agarwal and Kumar (2016).		Yes		Yes	Yes
9	Last and Eyal (2005)			Yes		Yes
10	Liao, Tsai and Su (2000)		Yes			
11	Maity and Maiti (2007).					Yes
12	Mehta and Shah (2003)		Yes			
13	Michalewicz (1992)					Yes
14	Pezzella, Morgantia and Ciaschettib (2008)					
15	Mishra (1979)		Yes			
16	Sarkar, Mukherjee and Balan (1997)		Yes		Yes	
17	Singh, Singh and Singh (2007)	Yes	Yes		Yes	
18	Singh, Kumari and Kumar (2010)	Yes		yes	Yes	
19	Singh and Singh (2010)	Yes	Yes			
20	Singh and Singh (2010)	Yes				
21	Teng and Chang (1995)		Yes		Yes	
22	Teng and Chang (2005)		Yes		Yes	
23	Wu, Ouyang and Yang (2006)				Yes	
24	Yao and Lee (1999)				Yes	
25	Yao, Chang and Su (2000)				Yes	

This paper presents two situations A: $S \leq S_0$ and B: $S > S_0$, where S_0 is the fixed stock level and S is the stock-level when production is terminated. Its major objective is to maximize the total profit on different situations and the environment, along with linear demand rate which is directly related to stock - level, if demand will increase or decrease then the stock level will also increase or decrease. Moreover, Centroid method is used to defuzzify the total profit and we can use GA to improve the result.

Genetic Algorithm (GA)

Genetic Algorithm is the method of computer engineering and operational research which is derived from Charles Darwin theory of biology. In GA, a population is modified until we will not get the optimized solution of our problem. GA uses four major functions to find the optimal solution which are:

(I) Fitness: It is determined by an objective or subjective function.
(II) Selection: After getting fitness population we select better population to perform further operations.
(III) Crossover: This is a function of combining two parents for next generation.
(IV) Mutation**:** It is a function of random variations in individual parents.

Where C_p is the probability of crossover, M_p is the probability of mutation, I_M is an iteration counter in each generation, I_{M0} is the upper limit of iteration counter I_M and I_N is the population size. L is the generation counter, Po (L) is the population of possible solutions for the generation L.

GA Algorithm:

1. Start and
2. Initializing the integer variables L, Po (L), I_M, Po_1 (L), C_p, M_p, I_{M0}, and I_N.
3. Set L = 0 and $I_M = 0$.
4. Evaluate the variable Po (L).
5. While ($I_M < I_{M0}$) otherwise go to step 16.
6. Select I_N solutions from Po (L) for mating pool using Roulette -Wheel process.
7. Select the solutions from Po (L), for crossover depending on C_p.
8. Now perform crossover operations on the selected solutions.
9. Select the solutions from Po (L), for mutation depending on M_p.
10. Perform mutation operation on selected solutions for getting the population Po_1 (L).
11. Classify the value of Po_1 (L).
12. Set $I_M = I_M + 1$.
13. If the average fitness of Po_1 (L) is greater than average fitness of Po (L) then go to steps 14,15,16. Otherwise, go to step 5
14. Set Po (L +1) = Po_1 (L).
15. Increment the value of L by L+1.
16. Set $I_M = 0$.
17. Print the Best solution value of Po (L).

GA Flow Chart:

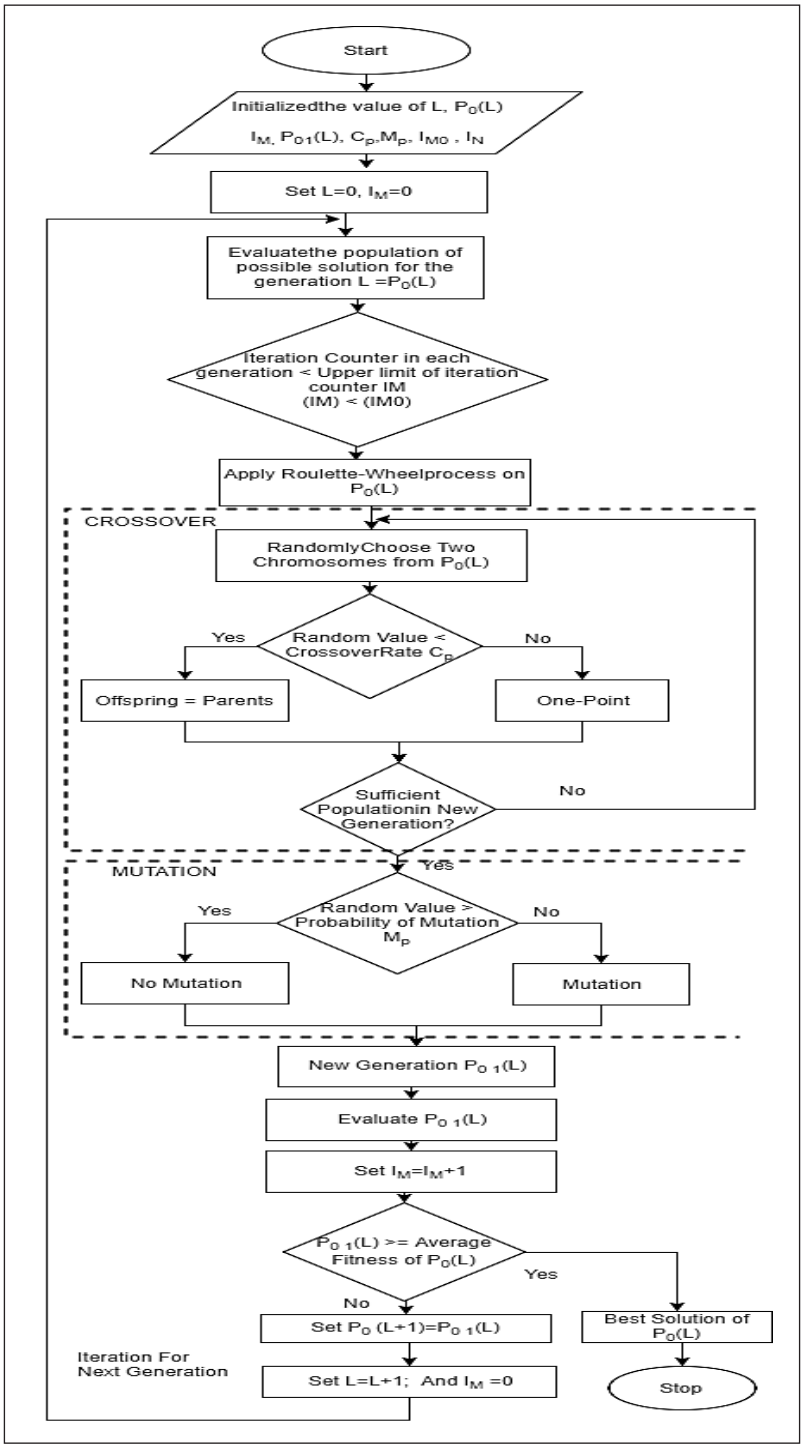


Figure 1. Genetic algorithm flow chart

Concept of Triangular Fuzzy Number

Let A= (k_1, k_2, k_3) is a triangular fuzzy number, where $k_1 = k - \Delta_1$, $k_2 = k$ and $k_3 = k + \Delta_2$. Such that $0 < \Delta_1 < k$, $0 < \Delta_2$ and Δ_1, Δ_2 are determined by the result producer depended on the uncertainty of the problem. And then A can be represented as $A = [A_l^L(\alpha), A_l^U(\alpha)] \cup [A_r^L(\alpha), A_r^U(\alpha)]$ subject to the constraint $0 < k_1 < k_2 < k_3$. And the membership function of the triangular fuzzy numbers is defined as follows. $\mu_A(x) : R \to [0,1]$

$$\mu_A(x) = \begin{cases} 0, x < k_1, x > k_3 \\ L(x) = \dfrac{x - k_1}{k_2 - k_1}, k_1 \le x \le k_2 \\ R(x) = \dfrac{k_3 - x}{k_3 - k_2}, k_2 \le x \le k_3 \end{cases}$$

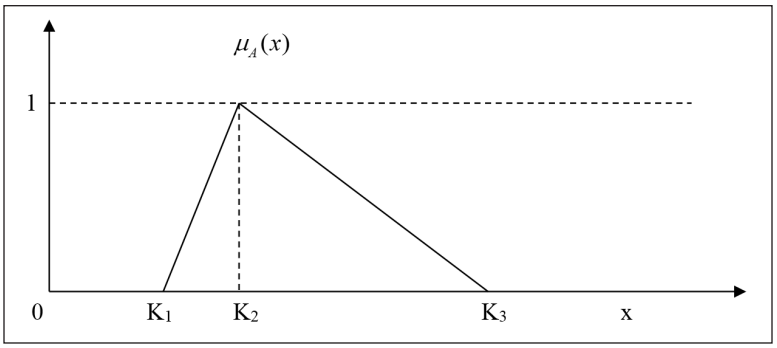


Figure 2. Graphical representation of membership function

Centroid Method

This method is also called the center of gravity or the center of the center. This is the most influential and appeals to all the defuzzification methods given by the arithmetic expression:

$$M_A = \frac{\int \mu_A(x) * x dx}{\int \mu_A(x) dx}$$

Where $\int$ denotes an algebraic integration. Hence the centroid for A is given as

$$M_k = \frac{K_1 + K_2 + K_3}{3} = \frac{k - \Delta_1 + k + k + \Delta_2}{3} = k + \frac{(\Delta_2 - \Delta_1)}{3}$$

Assumptions and Notations

The mathematical model of this study has been developed based on the following assumptions and notations.

Notations. We are using following notations for this model

$\tilde{c}_1$: Fuzzy holding cost of the inventory item, Rs./ per unit / per unit time.

$\tilde{c}_3$: Fuzzy deterioration cost, Rs. / per unit.

$\tilde{c}_3$: Fuzzy ordering cost, Rs. / per order.

$\tilde{c}_{1r}$: Fuzzy holding cost for raw material inventory, Rs. / per unit / per unit.

$\tilde{c}_{2r}$: Fuzzy deterioration cost for raw material inventory, Rs. / per unit.

$\tilde{c}_{3r}$: Fuzzy ordering cost for raw material inventory, Rs. / per order.

$\tilde{c}_r$: Fuzzy purchasing cost for raw material inventory, Rs./ per unit.

$\tilde{c}$: Fuzzy production cost, Rs. / per unit.

$\tilde{s}$: Fuzzy selling price, Rs. / per unit.

γ : Deterioration Rate.

I: Inflation Rate.
D: Discount Rate.
R: D-I
S: Maximum inventory level of the production cycle.
S_0 : Constant inventory level.
L: Length of the ordering cycle.
I (T): Inventory Level at time T ∈ [0, L].

Assumptions. The following assumptions, we are using in this model are:

1. There is only one item in the inventory system and the plan horizon is infinite.
2. The Production rate is demand dependent, P(T) = z D(T), where z is a non-negative constant.
3. The Demand rate is D(T) deterministic and given by its functional form

$$D(T) = \begin{cases} \alpha + \beta I(T), & S > S_0 \\ D, & 0 \le S \le S_0 \end{cases}$$

   Where α < 0, 0 < β < 1, D = c + b (T), and α, β, c, and d are known as parameters of scale and size, respectively.
4. Shortages are not permissible here.

Model Formulation

When the stock level reaches level S, S_0 has stopped production at constant stock-level. Then, in the following two cases, can arise

A: $S \leq S_0$

B: $S > S_0$

Case A: $S \leq S_0$. This is the standard Economic Production Quantity (EPQ) model for deteriorating goods with linear demand rate. At the time T = 0 output is started, with zero list level and at that time T=L_1 is stopped when the inventory level reaches level S. Due to the combined effect of the demand of that inventory level decreases and the fall in L, on which the inventory level reaches zero level. At any point of time the inventory level can be described by the following differential equation and graphically depicted in the figure-3.

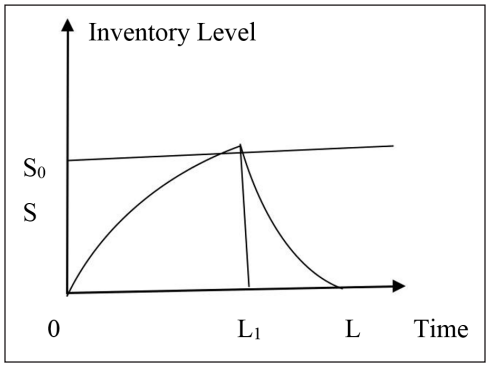


Figure 3. Graphical representation of case A: ($S \leq S_0$)

$$I_1'(T) + \gamma I_1(T) = (z-1)D, \ 0 \leq T \leq L_1 \tag{1}$$

$$I_2'(T) + \gamma I_2(T) = -D, \ L_1 \leq T \leq L \tag{2}$$

The boundary equations are $I_1(0) = 0$ and $I_2(L) = 0$. Solutions of the equation (1) and (2) are:

$$I_1(T) = \frac{(z-1)D}{\gamma} - \frac{(z-1)D}{\gamma} e^{(-\gamma T)} = \frac{(z-1)D}{\gamma}[1 - e^{(-\gamma T)}], \ 0 \leq T \leq L_1 \tag{3}$$

$$I_2(T) = \frac{De^{(\gamma L)}}{\gamma e^{(\gamma t)}} - \frac{D}{\gamma} = \frac{D}{\gamma}\left[e^{\gamma(L-T)} - 1\right], \ L_1 \leq T \leq L \tag{4}$$

From the equations of (3) and (4), Put $I_1(L_1) = I_2(L_1)$, then

$$e^{\gamma L} = ze^{\gamma L_1} - a + 1$$
$$\gamma L = \log[z(e^{\gamma L_1} - 1) + 1]$$
$$L = \frac{1}{\gamma}\log[z(e^{\gamma L_1} - 1) + 1] \tag{5}$$

The Maximum inventory level is $S = I_1(L_1) = \frac{(z-1)D}{\gamma}[1 - e^{(-\gamma L_1)}]$ (6)

The present value of the fuzzy holding cost of organized inventory is

$$HC = \frac{C_h D}{\gamma^2}\left\{(z-1)e^{(-\gamma L_1)}[(\gamma L_1 - 1)e^{\gamma L_1} + 1] + e^{(-\gamma L_1)}(e^{\gamma L} + \gamma L_1 e^{\gamma L_1}) - \gamma L + 1\right\} \tag{7}$$

The Present worth of fuzzy production cost is $PC = \frac{C_p D}{R}(1 - e^{(-RL_1)})$ (8)

The Present worth of fuzzy deteriorating cost is

$$DC = \frac{C_d D}{\gamma}\left\{(z-1)e^{(-\gamma L_1)}[(\gamma L_1 - 1)e^{\gamma L_1} + 1] + e^{(-\gamma L_1)}(e^{\gamma L} + \gamma L_1 e^{\gamma L_1}) - \gamma L + 1\right\} \tag{9}$$

The Present worth of fuzzy sales revenue is $SR = \frac{SD}{R}[1 - e^{(-RL)}]$ (10)

The Present worth of fuzzy ordering cost is $OC = C_3$ (11)

Raw Material Inventory Model during the Time (0, L_1)

Initially, the seller buys raw materials in many and produces finished goods. The producer starts production on time T = 0, and due to the combined effect of the production and decline of raw materials, T = L_1 reaches zero level in time. At any time, the seller's inventory system of raw material can be represented by the following differential equation and shown in the figure 4.

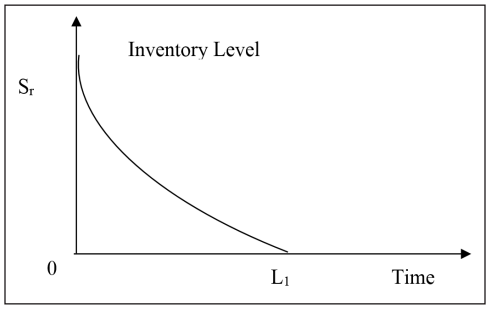


Figure 4. Graphical representation of raw material inventory

$$I_R'(T) + \gamma I_R(T) = -z(c + bT),\ 0 \le T \le L_1 \tag{12}$$

Solve the above equation with boundary condition I (L_1) = 0

$$I_R(T) = \frac{z}{\gamma^2}[e^{\gamma(L_1 - T)}(bL_1\gamma + c\gamma - b) - (\gamma bT + c\gamma - b)],\ 0 \le T \le L_1 \tag{13}$$

Maximum inventory level of the raw material, i.e. order quantity according to order from outer supplier

$$S_R = I_R(0) = \frac{z}{\gamma^2}\left[e^{L_1\gamma}(bL_1\gamma + c\gamma - b) - (c\gamma - b)\right] \tag{14}$$

Since the order is done at the time T = 0, thus the present worth of fuzzy ordering cost is

$$OC_R = A_R \tag{15}$$

The present worth of fuzzy holding cost for the raw material is

$$HC_R = \frac{h_R z}{\gamma^3}\left\{(e^{(\gamma L_1)} - 1)[bL_1\gamma + c\gamma - b] - \frac{\gamma L_1}{2}[bL_1\gamma + 2c\gamma - 2b]\right\} \tag{16}$$

The present worth of fuzzy deterioration cost for the raw material is

$$DC_R = \theta_R \frac{z}{\gamma^2}\left\{(e^{(\gamma L_1)} - 1)[bL_1\gamma + c\gamma - b] - \frac{\gamma L_1}{2}[bL_1\gamma + 2c\gamma - 2b]\right\} \tag{17}$$

Due to the cost of the item and the cost of the item being sold, the loss is incurred. Because the order is made on T = 0. The present value of the price of fuzzy items is

$$IC_R = P_R Q_R = P_R \frac{z}{\gamma^2}\left[e^{L_1\gamma}(bL_1\gamma + c\gamma - b) - (c\gamma - b)\right] \tag{18}$$

Therefore, the present worth of the total cost during the cycle is the cost of ordering costOC_R, holding cost HC_R, deterioration cost DC_R and item cost IC_R. Thus, for the raw materials, the present worth of total fuzzy cost is

$$TC_{1R} = OC_R + HC_R + DC_R + IC_R \tag{19}$$

Now, Present worth of total fuzzy profit is

$$TP_1 = SR - PC - HC - DC - OC - TC_{1R} \tag{20}$$

Observation 1. Total fuzzy profit (TP_{R1}) is the function of L_1. Now we have to find the optimal value of L_1 in order to maximize the total profit $TP_{R1}(L_1)$ subject to the inequality constraint $S \le S_0$, Now, we have LPP, **Max.** $TP_{R1}(L_1)$**; Subject to** $S_0 - S \ge 0$ (21)

Case B: $S \geq S_0$. In this case, the output starts with zero time, at time T = 0 and ends when the inventory level reaches level S, whereas T = L_2, where $\mathbf{S \geq S_0}$. Initially the demand and the production rate are constant up to the time T = L_1 at which inventory level reaches the level S_0 after that demand becomes stock dependent so as the production rate until the time T = L_2, after that inventory level is reduced because of the combined outcome of the demand and the deterioration spreads the zero level at the time T = L. The inventory level of this model can be explained by the following differential equations and graphically represented in Figure 5.

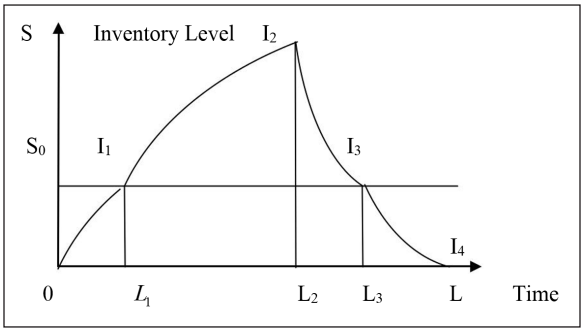


Figure 5. Graphical representation of case B: $S > S_0$

$$I_1'(T) + \gamma I_1(T) = (z-1)D, \ 0 \leq T \leq L_1 \tag{22}$$

$$I_2'(T) + \gamma I_2(T) = (z-1)(\alpha + \beta I_2(T)), \ L_1 \leq T \leq L_2 \tag{23}$$

$$I_3'(T) + \gamma I_3(T) = -[\alpha + \beta I_3(T)], \quad L_2 \leq T \leq L_3 \tag{24}$$

$$I_4'(T) + \gamma I_4(T) = -D, \ L_3 \leq T \leq L \tag{25}$$

Where the boundary limits are $I_1(0) = 0$, $I_2(L_1) = S_0$, $I_3(L_3) = S_0$, $I_4(L) = 0$.

Now solutions of above (22), (23), (24), (25) equations are:

$$I_1(T) = \frac{(z-1)D}{\gamma} - \frac{(z-1)D}{\gamma} e^{(-\gamma T)} = \frac{(z-1)D}{\gamma}[1 - e^{(-\gamma T)}], \ 0 \leq T \leq L_1 \tag{26}$$

$$I_2(T) = \frac{\alpha(z-1)}{(\gamma - z\beta + \beta)} + \left[S_0 - \frac{\alpha(z-1)}{(\gamma - z\beta + \beta)} \right] e^{(\gamma - z\beta + \beta)(L_1 - T)}, \ L_1 \leq T \leq L_2 \tag{27}$$

$$I_3(T) = -\frac{\alpha}{(\gamma + \beta)} + \left[S_0 + \frac{\alpha}{(\gamma + \beta)} \right] e^{(\gamma + \beta)(L_3 - T)}, \quad L_2 \leq T \leq L_3 \tag{28}$$

$$I_4(T) = \frac{De^{(\gamma L)}}{\gamma e^{(\gamma T)}} - \frac{D}{\gamma} = \frac{D}{\gamma} e^{\gamma(L-T)} - 1, \quad L_3 \le T \le L \tag{29}$$

Let I_1 $L_1 = S_0$, so equation (26) is

$$e^{\gamma L_1} = \frac{(z-1)D}{(z-1)D - S_0\gamma}, \quad L_1 = \frac{1}{\gamma}\log\left[\frac{(z-1)D}{(z-1)D - S_0\gamma}\right] \tag{30}$$

Now, equations from (27) and (28), Let $I_2 L_2 = I_3 L_2$, So

$$L_3 = L_2 + \frac{1}{\beta+\gamma}\log\left\{\frac{1}{S_0 + \frac{\alpha}{\beta+\gamma}}\left[\frac{\alpha(z-1)}{\beta+\gamma-z\beta} + \left(S_0 - \frac{\alpha(z-1)}{\beta+\gamma-z\beta}\right)e^{(\beta+\gamma-z\beta)(L_1-L_2)} + \frac{\alpha}{\beta+\gamma}\right]\right\} \tag{31}$$

Now equation from (29) Put condition $I_4 L_3 = S_0$

$$L = L_3 + \frac{1}{\gamma}\log\left[\frac{S_0 * \gamma}{D} + 1\right] \tag{32}$$

Value of L_3 put in equation (32), so

$$L = L_2 + \frac{1}{\beta+\gamma}\log\left\{\frac{1}{S_0 + \frac{\alpha}{\beta+\gamma}}\left[\frac{\alpha(z-1)}{\beta+\gamma-z\beta} + \left(S_0 - \frac{\alpha(z-1)}{\beta+\gamma-z\beta}\right)e^{(\beta+\gamma-z\beta)(L_1-L_2)} + \frac{\alpha}{\beta+\gamma}\right]\right\} + \frac{1}{\gamma}\log\left[\frac{S_0 * \gamma}{D} + 1\right] \tag{33}$$

So, maximum inventory level is $S = I_2(L_2)$

$$S = I_2(L_2) = \frac{\alpha(z-1)}{(\gamma - z\beta + \beta)} + \left[S_0 - \frac{\alpha(z-1)}{(\gamma - z\beta + \beta)}\right]e^{(\gamma - z\beta + \beta)(L_1 - L_2)} \tag{34}$$

Now, fuzzy holding cost is

$$\%\%\%\%\% = \%_1\left[\int_0^{L_1} I_1(T)dT + \int_{L_1}^{L_2} I_2(T)dT + \int_{L_2}^{L_3} I_3(T)dT + \int_{L_3}^{L} I_4(T)dT\right] \tag{35}$$

$$\%\%\%\%\% = \%_1\frac{D(z-1)}{\gamma}\left[L_1 - \frac{1}{\gamma} + \frac{e^{(-\gamma L_1)}}{\gamma}\right] +$$

$$\frac{\%_1}{(\gamma+\beta-z\beta)}\left[\alpha L_1(z-1) + \alpha L_2(z-1) + S_0 - S_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)} + \frac{\alpha(z-1)e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)} - \frac{\alpha(z-1)}{(\gamma+\beta-z\beta)}\right]$$

$$+ \frac{\%_1}{(\gamma+\beta)}\left[\alpha L_2 - \alpha L_3 + S_0 e^{(L_3-L_2)(\gamma+\beta)} - S_0 + \frac{\alpha}{(\beta+\gamma)}S_0 e^{(L_3-L_2)(\gamma+\beta)} - \frac{\alpha}{\beta+\gamma}\right] + \%_1\frac{D}{\gamma}\left[L_3 - \frac{1}{\gamma} - L + \frac{e^{\gamma(L-L_3)}}{\gamma}\right] \tag{36}$$

fuzzy deterioration cost is

$$DC = \theta\gamma\left[\int_0^{L_1} I_1(T)dT + \int_{L_1}^{L_2} I_2(T)dT + \int_{L_2}^{L_3} I_3(T)dT + \int_{L_3}^{L} I_4(T)dT\right]$$

$$DC = \theta D(z-1)\left[L_1 - \frac{1}{\gamma} + \frac{e^{(-\gamma L_1)}}{\gamma}\right] +$$

$$\frac{\theta\gamma}{(\gamma+\beta-z\beta)}\left[\alpha L_1(z-1) + \alpha L_2(z-1) + S_0 - S_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)} + \frac{\alpha(z-1)e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)} - \frac{\alpha(z-1)}{(\gamma+\beta-z\beta)}\right]$$

$$+\frac{\theta\gamma}{(\gamma+\beta)}\left[\alpha L_2 - \alpha L_3 + S_0 e^{(L_3-L_2)(\gamma+\beta)} - S_0 + \frac{\alpha}{(\beta+\gamma)} S_0 e^{(L_3-L_2)(\gamma+\beta)} - \frac{\alpha}{\beta+\gamma}\right] + \theta D\left[L_3 - \frac{1}{\gamma} - L + \frac{e^{\gamma(L-L_3)}}{\gamma}\right]$$

(37)

fuzzy production cost is

$$PC = pD(z-1)\left[L_1 - \frac{1}{\gamma} + \frac{e^{(-\gamma L_1)}}{\gamma}\right] +$$

$$\frac{p\gamma}{(\gamma+\beta-z\beta)}\left[\alpha L_1(z-1) + \alpha L_2(z-1) + S_0 - S_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)} + \frac{\alpha(z-1)e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)} - \frac{\alpha(z-1)}{(\gamma+\beta-z\beta)}\right]$$

(38)

fuzzy sales revenue is

$$SR = \frac{sD(z-1)}{\gamma}\left[L_1 - \frac{1}{\gamma} + \frac{e^{(-\gamma L_1)}}{\gamma}\right] +$$

$$\frac{s}{(\gamma+\beta-z\beta)}\left[\alpha L_1(z-1) + \alpha L_2(z-1) + S_0 - S_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)} + \frac{\alpha(z-1)e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)} - \frac{\alpha(z-1)}{(\gamma+\beta-z\beta)}\right]$$

$$+\frac{s}{(\gamma+\beta)}\left[\alpha L_2 - \alpha L_3 + S_0 e^{(L_3-L_2)(\gamma+\beta)} - S_0 + \frac{\alpha}{(\beta+\gamma)} S_0 e^{(L_3-L_2)(\gamma+\beta)} - \frac{\alpha}{\beta+\gamma}\right] + \frac{sD}{\gamma}\left[L_3 - \frac{1}{\gamma} - L + \frac{e^{\gamma(L-L_3)}}{\gamma}\right]$$

(39)

fuzzy ordering cost is $OC = A$ (40)

Raw Material Inventory Model during the Time (0, L)

Originally, the vendor purchased the raw material in lots and then produces the finalized material. The Vendor starts production at time Zero (T = 0), and the raw material reaches the zero level at the time (T = L_2) because of the collective effect of production and

deterioration. The vendor's raw materials inventory system at any time T can be denoted by the following differential equation and has shown in Figure 6.

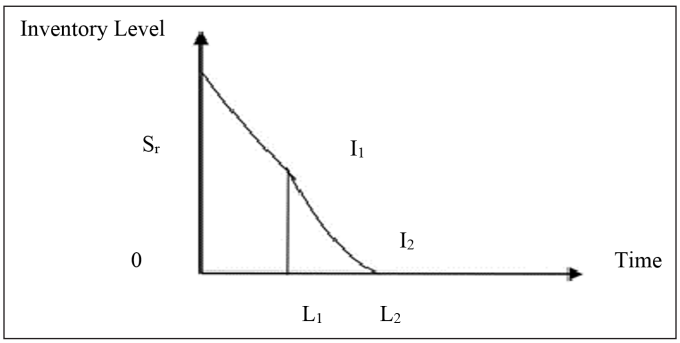


Figure 6. Graphical representation of raw material inventory

$$I_{R1}'(T)+\gamma I_{R1}(T)=-z(c+bT),\ 0\le T\le L_1 \tag{41}$$

$$I_{R2}'(T)+\ I_{R2}(T)=-z[\alpha+\beta I_2(T)],\ L_1\le T\le L_2 \tag{42}$$

The value of $I_2(T)=\dfrac{\alpha(z-1)}{(\gamma-z\beta+\beta)}+\left[S_0-\dfrac{\alpha(z-1)}{(\gamma-z\beta+\beta)}\right]e^{(\gamma-z\beta+\beta)(L_1-T)}$ from the equation (27), So

$$I_{R2}'(T)+\gamma I_{R2}(T)=-z\left\{\alpha+\beta\left\{\frac{\alpha(z-1)}{(\gamma-z\beta+\beta)}+\left[S_0-\frac{\alpha(z-1)}{(\gamma-z\beta+\beta)}\right]e^{(\gamma-z\beta+\beta)(L_1-T)}\right\}\right\} \tag{43}$$

Where the boundary conditions are $I_{R1}(0)=S_R$ and $I_{R2}(L_2)=0$,

Now Solution of above equations (41) and (43) are

$$I_{R1}(T)=-zT(c+bT)+S_Re^{(-T\gamma)},\ 0\le T\le L_1 \tag{44}$$

$$I_{R2}(T)=-\frac{z\alpha}{\gamma}-\frac{z(z-1)\beta\alpha}{\gamma(\gamma+\beta-z\beta)}-\frac{zS_0e^{(\gamma+\beta-z\beta)(L_1-T)}}{(z-1)}+\frac{z\alpha e^{(\gamma+\beta-z\beta)(L_1-T)}}{(\gamma+\beta-z\beta)}+\frac{z\alpha e^{\gamma(L_2-T)}}{\gamma}$$
$$\frac{z(z-1)\alpha\beta e^{\gamma(L_2-T)}}{\gamma(\gamma+\beta-z\beta)}+\frac{zS_0e^{(\gamma+\beta-z\beta)(L_1-L_2)+(\gamma L_2)}}{(z-1)e^{T\gamma}}-\frac{z\alpha e^{(\gamma+\beta-z\beta)(L_1-L_2)+(\gamma L_2)-(\gamma T)}}{(\gamma+\beta-z\beta)},\ L_1\le T\le L_2 \tag{45}$$

Now using the condition $I_{R1}(L_1)=I_{R2}(L_1)$, The order quantity per order from outside suppliers. The maximum inventory level of the raw materials.

$$S_R = zL_1(c+bL_1)e^{\gamma L_1} - \frac{z\alpha e^{\gamma L_1}}{\gamma} - \frac{e^{\gamma L_1} z(z-1)\beta\alpha}{z(\gamma+\beta-z\beta)} - \frac{ze^{\gamma L_1} S_0}{(z-1)} + \frac{e^{\gamma L_1} z\alpha}{\gamma+\beta-\beta z} + \frac{z\alpha e^{\gamma L_2}}{\gamma} + \frac{z(z-1)\beta\alpha e^{\gamma L_2}}{\gamma(\gamma+\beta-z\beta)} + \frac{zS_0 e^{\gamma L_2} e^{(L_1-L_2)(\beta+\gamma-z\beta)}}{(z-1)} - \frac{z\alpha e^{\gamma L_2} e^{(L_1-L_2)(\beta+\gamma-z\beta)}}{(\gamma+\beta-z\beta)} \tag{46}$$

From the order is done at the time (T= 0), then the fuzzy ordering cost is

$$OC = A \tag{47}$$

Now, Fuzzy holding cost is $HC_R = h\left[\int_0^{L_1} I_{R1}(T)dT + \int_{L_1}^{L_2} I_{R2}(T)dT\right]$

$$HC_R = h\left[\begin{array}{l} S_R\left(\frac{1}{\gamma} - \frac{e^{-\gamma L_1}}{\gamma}\right) - \frac{z(2bL_1^3 + 3cL_1^2)}{6} - \frac{z\alpha L_2}{\gamma} + \frac{z\alpha L_1}{\gamma} + \frac{z(z-1)\beta\alpha L_1}{\gamma(\gamma+\beta-z\beta)} \\ - \frac{z(z-1)\beta\alpha L_2}{\gamma(\gamma+\beta-z\beta)} + \frac{zS_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)(z-1)} - \frac{zS_0}{(\gamma+\beta-z\beta)(z-1)} - \frac{z\alpha e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)^2} \\ + \frac{z\alpha}{(\gamma+\beta-z\beta)^2} + \frac{z\alpha e^{(L_2-L_1)}}{\gamma^2} - \frac{z\alpha}{\gamma^2} - \frac{z\beta\alpha(z-1)}{(\gamma+\beta-z\beta)\gamma^2} + \frac{\alpha\beta z(z-1)e^{\gamma(L_2-L_1)}}{(\gamma+\beta-z\beta)\gamma^2} \\ + \frac{zS_0 e^{(-z\beta L_1 + z\beta L_2 - \beta L_2 + \beta L_1)}}{\gamma(z-1)} - \frac{zS_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{\gamma(z-1)} + \frac{z\alpha e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)\gamma} - \frac{z\alpha e^{\gamma(L_2-L_1)} e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)(\gamma)} \end{array}\right] \tag{48}$$

So, Fuzzy deterioration cost is

$$DC_R = h\gamma\left[\int_0^{L_1} I_{R1}(T)dT + \int_{L_1}^{L_2} I_{R2}(T)dT\right]$$

$$DC_R = h\gamma\left[\begin{array}{l} S_R\left(\frac{1}{\gamma} - \frac{e^{-\gamma L_1}}{\gamma}\right) - \frac{z(2bL_1^3 + 3cL_1^2)}{6} - \frac{z\alpha L_2}{\gamma} + \frac{z\alpha L_1}{\gamma} + \frac{z(z-1)\beta\alpha L_1}{\gamma(\gamma+\beta-z\beta)} \\ - \frac{z(z-1)\beta\alpha L_2}{\gamma(\gamma+\beta-z\beta)} + \frac{zS_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)(z-1)} - \frac{zS_0}{(\gamma+\beta-z\beta)(z-1)} - \frac{z\alpha e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)^2} \\ + \frac{z\alpha}{(\gamma+\beta-z\beta)^2} + \frac{z\alpha e^{(L_2-L_1)}}{\gamma^2} - \frac{z\alpha}{\gamma^2} - \frac{z\beta\alpha(z-1)}{(\gamma+\beta-z\beta)\gamma^2} + \frac{\alpha\beta z(z-1)e^{\gamma(L_2-L_1)}}{(\gamma+\beta-z\beta)\gamma^2} \\ + \frac{zS_0 e^{(-z\beta L_1 + z\beta L_2 - \beta L_2 + \beta L_1)}}{\gamma(z-1)} - \frac{zS_0 e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{\gamma(z-1)} + \frac{z\alpha e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)\gamma} - \frac{z\alpha e^{\gamma(L_2-L_1)} e^{(L_1-L_2)(\gamma+\beta-z\beta)}}{(\gamma+\beta-z\beta)(\gamma)} \end{array}\right] \tag{49}$$

Now the loss of new item costs and the cost of the item sold is also included. Because the order is done at T = 0. And the cost of fuzzy items

$$IoC_R = I_R S_R \tag{50}$$

So, the total cost during the cycle is the sum of ordering cost OoC_R, holding cost HoC_R, deterioration cost DoC_R and fuzzy item cost IoC_R. Thus, for the raw materials, the present worth of total fuzzy cost is

$$(TP)_{2R} = OoC_R + HoC_R + DoC_R + IoC_R \tag{51}$$

Now our the Present worth of total fuzzy profit is

$$TP_{R2} = SoR - PoC - HoC - DoC - OoC - IoC_{2R} \tag{52}$$

Observation 2. The total fuzzy profit (TP_{R2}) is the function of L_2. Now we have to find the optimum value of L_2 to maximize the total profit $TP_{R2}(L_2)$ subject to the inequality constraint $S > S_0$, Now, we have LPP, **Max.** TP_{R2} **(L_2); Subject to S - S_0 > 0** (53)

Numerical Example

Case A: $S_0 \geq S$. The following values are used to interpret the model.

z = 1.2, b = 1.4, c=1.5, α=25, β=0, S_0 = 150, I_1= 0.45, 0.50, 0.60, I_2= 0.35, 0.40, 0.50, I_3 =130, 200, 300, I_{R1} = 0.45, 0.50, 0.60, I_{R2} = 0.35,0.40, 0.50, I_{3R} = 80, 100, 125, I_R = 0.8,1.0,1.2

β_c = 4,5,7, I_c= 10, 12, 15, D = 0.1, I = 0.05, γ =0.01.

From the above parametric values, we get the results with the help of GA and the results are shown in Table 2.

Table 2
Results of optimal values

L_1	L	S_r	S	TP_{r1} (L_1)
4.14	4.95	47.93	5.84	960.46

We receive the optimum cyclic length L= 4.95, and the maximum inventory level S = 5.84 units and the output time is L_1= 4.14, the optimum gain is 960.46.

At the beginning of each cycle, the raw material of S_R = 47.93 units for the L_1 Purchases producer of P and produces the final product. In this time period, the raw material list decreases and on time T = 4.14, it will reach zero level, the list of final products will be build up time T = 4.14 and the inventory level will reach 5.84 units. At this time, production and

inventory ends due to the impact of manufacturer demand and deterioration. The inventory level reaches zero at the time level T = 4.95, then the output will start for the next cycle.

Sensitivity analysis. Effect of different parameters on Total Profit.

Table 3
Present value of total profits of the model due to z

S. No.	z	TP_{r1}
1.	1.3	969.24
2.	1.4	972.45
3.	1.5	958.41
4.	1.6	942.25

Table 4
Present value of total profits of the model due to α

S. No.	α	TP_{r1}
1.	20	756.12
2.	26	975.15
3.	27	985.48
4.	29	994.15

Table 5
Present value of total profits of the model due to β

S. No.	β	TP_{r1}
1.	0.23	742.15
2.	0.27	969.49
3.	0.29	987.81
4.	0.30	996.17

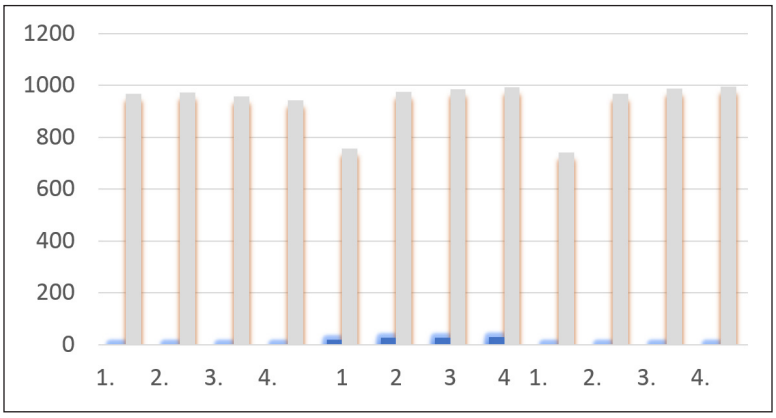


Figure 7. Graphical representation of different parameters z, α, β and their Total Profit TP_{r1}

Observations.

1. From the above table number 3, 4, and 5 with different parameters, we look at the different values of total profit in relation to different production rates.
2. We have seen that if the production rate increases, then the total profit also gets the maximum increase in production but the profit will be reduced.
3. Due to the linear demand rate stock level is increasing and the production rate also increases.
4. If the production rate increases by fixed limit then decrease in optimum profit due to the increase in the cost of the fall.
5. The results of different values of demand parameters are obtained in table 4.
6. We have demanded in reality that the optimum profit increases if demand criteria are increasing.
7. With the effect of inflation and discount rate, the results of parametric values are present in Table 5.
8. We have seen that if the effect of inflation and rebate rate is on the rise then optimal profit is reduced.

Case B: $S_0 < S$. We are using the following values to describe the model.

z = 1.2, b = 1.4, c =1.5, α=25, β=0.25, S_0 = 150, θ_1= 0.45, 0.50, 0.60, θ_2= 0.35, 0.40, 0.50,

θ_3=130, 200, 300, θ_{R1} = 0.45, 0.50, 0.60, θ_{R2} = 0.35,0.40, 0.50, θ_{3R} = 80, 100, 125, θ_R = 0.8,1.0,1.2

β_c = 4,5,7, θ_c= 10, 12, 15, D = 0.1, I = 0.05, γ =0.01, R = 0.05.

With the above parametric values, we get results with the help of GA and the results are in Table 6.

We receive optimum cyclic length L= 87.14, and maximum inventory level S = 232.5 units and the production time L_1= 46.20, the optimum profit is 62348.154.

At the beginning of each cycle, S_r =847.481 units of the raw materials for the period L_1 purchases, a manufacturer of rate P and production the final product. In this time period, the raw material list decreases and at time T = 52.67, it will reach zero level, the inventory of final product will be T = 52.67 and the inventory level will reach 232.5 units. At this time, production and inventory ends due to the impact of manufacturer demand and fall. Inventory level reaches zero level at time T = 87.14, then production will start for next cycle.

Table 6
Results of optimal values

L_1	L_2	L_3	L	S_r	S	T P_{r2} (L_2)
46.20	52.67	45.27	87.14	847.481	232.5	62348.154

Sensitivity analysis. Effect of various parameters on Total Profit.

Table 7
Present value of total profits of the model due to z

S. No.	z	TP_{r2}
1.	1.25	94454.15
2.	1.30	135785.54
3.	1.35	945175.29
4.	1.40	1245475.595

Table 8
Present value of total profits of the model due to α

S. No.	α	TP_{r2}
1.	24	48972.28
2.	26	69547.53
3.	27	84981.34
4.	28	94724.87

Table 9
Present value of total profits of the model due to β

S. No.	β	TP_{r2}
	0.24	57415.26
	0.26	78445.58
	0.27	81156.24
	0.28	94584.15

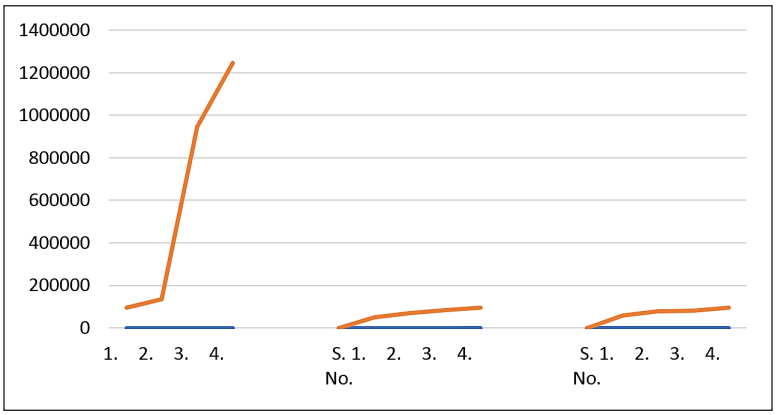


Figure 8. Graphical representation of different parameters z, α, β and their Total Profit TP_{r2}

Observations.

1. From the above table number 7, 8, and 9 with different parameters, we look at the different values of total profit in relation to different production rates.
2. We have seen that if the production rate increases, then the total profit also gets the maximum increase in production but the profit will be reduced.
3. Due to the linear demand rate stock level is increasing and the production rate also increases.
4. If the production rate exceeds a certain limit, then the reduction in optimum profit due to increase in the cost of the cost of the holding and the cost of the fall.
5. Results in the table 8 are obtained for different standards of demand parameters.
6. We have demanded in reality that the optimum profit increases if demand criteria are increasing.
7. With the effect of inflation and discount rate, the results of parametric values are present in Table 9.
8. We have seen that if inflation and rebate rate are increasing, then optimum profit is also decreasing.

CONCLUSIONS AND FUTURE WORK

A production inventory model developed for newly launched product with the effect of inflation under fuzzy environment. And the demand rate is considered linearly stock dependent and shortage are not allowed. Model is formulated to maximize the expected profit from the full planning horizon. Also, we used to Genetic Algorithm (GA) for optimal extreme solutions for growth of business. A Genetic Algorithm with varying population size is used to solve the model where crossover probability is a function of parent's age type and it is obtained by fuzzy rule-base and possibility theory. And next section, the model is developed under the effect of inflation in fuzzy environment. This GA can also help in resolving the decision-making problems in the areas of science and technology and forecasting.

This fuzzy inventory model can be extended to supermarket price dependent demand, GDP policy, to storage and cold storage industry.

In this study, we develop an EPQ model with stock depended demand rate under the influence of inflation in a fuzzy environment, for the defuzzification we used centroid method and Genetic Algorithm (GA) is used for optimization.

So far, no researchers have considered these factors in their studies.

REFERENCES

Agarwal, P., Sharma, A., & Kumar, N. (2017). Inventory control in ice and cold storage management using Genetic Algorithm. *International Journal of Computer & Mathematical Sciences, 6*(6), 2347-8527.

Balkhi, T. Z., & Benkherouf, L. (2004). On an inventory model for deteriorating items with stock dependent and time varying demand rates. *Computers and Operations Research, 31*(1), 223-240.

Bessaou, M., & Siarry, P. (2001). A genetic algorithm with real-value coding to optimize multimodal continuous function. *Structural and Multi-Disciplinary Optimization, 23*(1), 63-74.

Buzacott, J. A. (1975). Economic order quantities with inflation. *Operation Research Quarterly, 26*(1), 553-558.

Chandra, M. J., & Bahner, M. J. (1985). The effects of inflation and time value of money on some inventory systems. *International Journal of Production Research, 23*(1), 723-729.

Chang, H. C., Yao, J. S., & Quyang, L. Y. (2004). Fuzzy mixture inventory model with variable lead-time based on probabilistic fuzzy set and triangular fuzzy number. *Computer and Mathematical Modeling, 39*(1), 287-304.

Datta, T. K., Paul, K., & Pal, A. K. (1998). Demand promotion by up-gradation under stock-dependent demand situation – a model. *International Journal of Production Economics, 55*(1), 31–38.

Kumar, S., Agarwal, P., & Kumar, N. (2016). An inventory model for deteriorating items under inflation and permissible delay in payments by genetic algorithm. *Cogent Business & Management, 3*(1), 53 – 68.

Last, M., & Eyal, S. (2005). A fuzzy-based lifetime extension of genetic algorithms. *Fuzzy Sets and Systems, 149*(1), 131-147.

Liao, H. C., Tsai, C. H., & Su, C. T. (2000). An inventory model with deteriorating items under inflation when a delay in payment is permissible. *International Journal of Production Economics, 63*(1), 207-214.

Maity, M. K., & Maiti, M. (2007). Two-storage inventory model with lot-size dependent fuzzy lead-time under possibility constraints via genetic algorithm. *European Journal of Operational Research, 179*(1), 352-371.

Mehta, N. J., & Shah, N. H. (2003). An inventory model for deteriorating items with exponentially increasing demand and shortages under inflation and time discounting. *Investigacao Operational, 23*(1), 103-111.

Michalewicz, Z. (1992). Genetic algorithms + data structures = evolution programs. *Mathematical Intelligencer, 18*(3), 71.

Mishra, R. B. (1979). A note on optimal inventory management under inflation. *Naval Research Logistics, 26*(1), 161-165.

Pezzella, F., Morgantia, G., & Ciaschettib, G. (2008). A genetic algorithm for the flexible job-shop scheduling problem. *Computers and Operations Research, 35*(1), 3202-3212.

Sarkar, B. R., Mukherjee, S., & Balan, C. V. (1997). An order level lot-size inventory model with inventory-level independent demand and deterioration. *International Journal of Production Economics, 48*(1), 227-236.

Singh, S. R., & Singh, C. (2010). Supply Chain Model with Stochastic Lead Time under Imprecise Partially Backlogging and Fuzzy Ramp-Type Demand for Expiring Items. *International Journal of Operational Research, 8*(4), 511-522.

Singh, S. R., & Singh, C. (2010). Two echelon supply chain model with imperfect production, for Weibull distribution deteriorating items under imprecise and inflationary environment. *International Journal of Operations Research and Optimization, 1*(1), 9-25.

Singh, S. R., Kumari, R., & Kumar, N. (2010). Replenishment policy for non-instantaneous deteriorating items with stock-dependent demand and partial back logging with two-storage facility under inflation. *International Journal of Operations Research and Optimization, 1*(1), 161-179.

Singh, S. R., Singh, C., & Singh, T. J. (2007). Optimal policy for decaying items with stock-dependent demand under inflation in a supply chain. *International Review of Pure and Applied Mathematics, 3*(2), 189-197.

Teng, J. T., & Chang, C. T. (1995). EPQ model for deteriorating items with price and stock-dependent demand. *Computers and Operations Research, 31*(1), 297-308.

Teng, J. T., & Chang, C. T. (2005). Economic production quantity models for deteriorating items with price and stock dependent demand. *Computers and Operational Research, 32*(1), 297-308.

Wu, K. S., Ouyang, L. Y., & Yang, C. T. (2006). An optimal replenishment policy for non-instantaneous deteriorating items with stock-dependent demand and partial backlogging. *International Journal of Production Economics, 101*(1), 369-384.

Yao, J. S., & Lee, H. M. (1999). Fuzzy inventory with or without backorder for fuzzy order quantity with trapezoidal fuzzy number, *Fuzzy Sets and Systems, 105*(1), 311-337.

Yao, J. S., Chang, S. C., & Su, J. S. (2000). Fuzzy inventory without backorder for fuzzy order quantity and fuzzy total demand quantity. *European Journal of Operational Research, 148*(1), 401-409.

Pertanika J. Sci. & Technol. 26 (4): 1659 - 1675 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Survey of Hand Gesture Recognition Methods in Sign Language Recognition

Suharjito[1]*, Meita Chandra Ariesta[2], Fanny Wiryana[2] and Gede Putra Kusuma[1]

[1]Computer Science Department, Bina Nusantara University (BINUS), Jakarta 11480, Indonesia,
[2]Computer Science Department, School of Computer Science, Bina Nusantara University (BINUS), Jakarta 11480, Indonesia

ABSTRACT

Sign Language is the only method used in communication between the hearing-impaired community and common community. Sign Language Recognition (SLR) system, which is required to recognize sign languages, has been widely studied for years. The studies are based on various input sensors, gesture segmentation, extraction of features and classification methods. This paper aims to analyze and compare the methods employed in the SLR systems, classifications methods that have been used, and suggests the most promising method for future research. Due to recent advancement in classification methods, many of the recent proposed works mainly contribute on the classification methods, such as hybrid method and Deep Learning. This paper focuses on the classification methods used in prior Sign Language Recognition system. Based on our review, HMM-based approaches have been explored extensively in prior research, including its modifications. Deep Learning such as Convolutional Neural Network is popular in the past five years. Hybrid CNN-HMM and fully Deep Learning approaches have shown promising results and offer opportunities for further exploration. However, overfitting and high computational requirements still hinder their adoption. We believe the future direction of the research is toward developing a simpler network that can achieve high performance and requires low computational load, which embeds the feature learner into the classifier in multi-layered neural network fashion.

ARTICLE INFO

Article history:
Received: 15 September 2017
Accepted: 17 April 2018
Published: 24 October 2018

E-mail addresses:
harjito@yahoo.com (Suharjito)
keroaduh@gmail.com (Meita Chandra Ariesta)
fannywirya@gmail.com (Fanny Wiryana)
inegara@binus.edu (Gede Putra Kusuma)
* Corresponding author

Keywords: Classification methods, Hidden Markov model, neural networks, sign language recognition

INTRODUCTION

Sign language is delivered through visual communications such as gestures, hand-shapes, facial expressions, and movement

ISSN: 0128-7680
e-ISSN: 2231-8526

of hands. Different from oral languages, sign language has its own vocabulary comprising combinations of various visual features for conveying messages in words or sentences (Sahoo, Mishra, & Ravulakollu, 2014). The communication between hearing impaired community and common community is limited mostly because of the common community's lack of knowledge about sign language. Thus, it is very hard for members of the common community to freely converse with hearing-impaired persons. This topic is very important as we are experiencing globalization and every person should receive the same opportunities regardless of their backgrounds and physical conditions. Therefore, a system which can translate sign language into common language and vice versa is needed. The system may help normal persons understand sign language so that the communication between hearing impaired and common communities can be easier.

This paper aims to analyze and compare the methods implemented in previous researches. Moreover, it aims to suggest the best method to explore for future research. We hope to create an Indonesian Sign Language Recognition system using the method suggested by this study. Several studies have been reported that review prior works in order to suggest the best method in Sign Language Recognition system. Majid and Zain (2013) reviewed the development of Sign Language Recognition system for different sign languages. They reviewed only 32 related publications up to year 2012. The review was focused on data acquisition and recognition methods. They suggested a sign language recognition system using hybrid Fuzzy and Neural Network with Kinect to tackle accuracy and efficiency problems. A comprehensive review on hand gesture recognition was reported by Pishardy et al. (Pisharady & Saerbeck, 2015). They reviewed 159 publications up to year 2015. The review paper contained prior works related to data acquisition, feature extraction, and classification methods. They also reviewed some hand gesture databases. They suggested that Time Delay Neural Network was compact and efficient to use because it optimized features detection and reduced training time, while Hidden Markov Model requires a large number of training data and high computational costs (Pisharady & Saerbeck, 2015). However, none the classification methods reviewed in their paper is about deep learning approach, which is recently getting popularity as a powerful classification method. In this work, we add some latest developments in the field including deep learning approaches, and provide more coverage on the classification methods. We reviewed 70 publications, among which 19 publications were from the year 2016 and 2017. These publications were obtained through Google Scholar search engine. We notice that many of the recent proposed works mainly contribute on the classification methods, such as hybrid method and deep learning, rather than on stages prior to the classification. Therefore, we focus our review on the classification methods. For completeness, we also report an overview of the pipeline of hand gesture recognition that consists of data acquisition, hand segmentation, feature extraction, and recognition methods in the next section.

Feature extraction, recognition methods, and implementation are essentials in the development of Sign Language Recognition (SLR) system. According to Bhuyan, Kumar, MacDorman and Iwahori (2014), determining the start and end point of a meaningful gesture, also called gesture segmentation, was one of many challenges in hand gesture recognition. Feature extraction is another problem. Zhang et al. optimized feature extraction process by detecting pupil and using it as a reference point as well as using colored-gloves for hand detection (Zhang, Chen, Fang, Chen, & Gao, 2004). In recent studies, Kinect had been widely used for image acquisitions in SLR system because it could track hand and body actions easily and accurately. It also provided depth and color data at the same time (Chai et al., 2013). On top of that, research on the classification method was also elaborated. Although feature extraction is important, the classification method employed is also important and still needs improvements. To overcome the difficulties in hand segmentation, hand tracking, and complex backgrounds in SLR systems, Huang, Zhou, Li and Li (2015) et al. implemented 3D Convolutional Neural Network to address these problems. There are many studies that experimented on various classification methods in SLR system. The methods used in the previous studies will be analyzed and compared further in this paper to find the best classification method, which will be used in an Indonesian Sign Language Recognition system in the future.

Hand Gesture Recognition

Ahuja and Singh (2015) explained the basic module of Hand Gesture Recognition. The basic module of Hand Gesture Recognition consists of four steps: Image Acquisition, Hand Segmentation, Feature Extraction, and Hand Gesture Recognition, as shown in Figure 1. Image acquisition is the process of capturing images for vision-based approach. Leap Motion Controller, Microsoft Kinect, Data Glove, and Vision-based are some other methods of sign acquisition. Leap Motion Controller (LMC) is a device that can detect hand movement up to 200 frames per second and assign ID to each detection (Mohandes, Aliyu, & Deriche, 2014). LMC converts signals into computer commands (Bhavsar, 2017; Kakde, Nakrani, & Rawate, 2016). LMC has been commonly used in prior studies related to hand gesture recognition (Koul, Patil, Nandurkar, & Patil, 2016; Potter, Araullo, & Carter, 2013). Microsoft Kinect is used to capture every motion and turns it into a feature by using the built-in 3D sensory camera. Microsoft Kinect is widely used in hand gesture recognition (Carneiro et al., 2016; Escobedo & Camara, 2016; Jiang, Tao, Ye, Wang, & Ye, 2014; Keskin, Kırac, Kara, & Akarun, 2011; Raheja, Mishra, & Chaudhary, 2016; Rakun, Andriani, Wiprayoga, Danniswara, & Tjandra, 2013). Data-Glove Based Method is a relatively old data acquisition method for gesture recognition. This method utilizes a device to help the process of collecting data. The device is a glove which has sensors

connected to a computer. Those sensors detect the movement and changes in the hands and fingers of the users (Mehdi, 2002; Phi, Nguyen, Bui, & Vu, 2015; Ranjini & M, 2014; Saengsri, Niennattrakul, & Ratanamahatana, 2012).

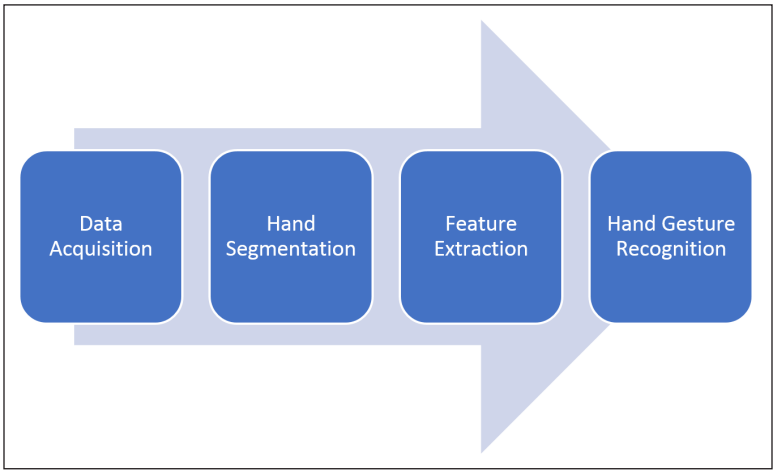


Figure 1. Basic hand gesture recognition module

Hand segmentation is a process to isolate hands and other features from the rest of the image in vision based systems. Zhang et al. (2004) employed pupil-detection algorithm to make the pupil as reference points and used colored gloves to assist the background segmentation. Many hand segmentation methods have been proposed in computer vision. Canny Edge Detector is used to detect the edges of hands from an image. Canny Edge Detector is known for its optimal performance in detecting edges and low error rate (Ghosh & Ari, 2016; Kalsh & Garewal, 2013; M. V. D. Prasad, Kishore, Kumar, & Kumar, 2016). The other method for hand segmentation is Elliptical Fourier Descriptors which is specialized for extracting outline of shapes (Kishore, Prasad, Prasad, & Rahul, 2015). Skin detection is also used in hand segmentation. This method simply identifies and segments the skin area from the rest of the image (Lim, Tan, & Tan, 2016; Ong & Ranganath, 2005). Pankajakshan & Thilagavathi (2015) also applied hand motion tracking with skin detection to produce more precise result Similar to skin detection, colored-gloves were used to give the hands a distictive feature, therefore assisting the hand segmentation process (Eugene Starner, 1995; Wang & Popović, 2009; Zhang, Chen, Fang, Chen, & Gao, 2004).

Feature extraction is used to acquire features from the images captured. The features include background data, translation, scale, shape, rotation, angle, coordinates, and movements (Bhavsar & Trivedi, n.d.). Yan et al. (2017) proposed multiscale Convolutional Neural Network to detect hands to tackle problems in prior studies regarding hand detection, which struggled due to low resolution, varying lighting intensity, heavy and occlusion. Scale

Invariant Feature Transform (SIFT) algorithm had been used as feature extraction methods in prior studies (Goyal & Singh, 2014; S. Goyal, Sharma, & Sharma, 2013). According to Gurjal and Kunmur (2012), SIFT algorithm was robust against rotation, translation, or scaling variation, and produced a large collection of local feature vectors.

Classification is the final stage and an important stage in gesture recognition (Ghosh & Ari, 2016). Word or sentence in sign language is made up of continuous gesture, which changes over time. Therefore, a recognition method should be able to handle sequential data. Some problems occur when the system handles noisy data and uncontrolled environment. Therefore, Roh et al. (2016) proposed Selective Temporal Filtering to tackle those problems. There are two types of Gesture Recognition approaches. Some researchers used the extracted features for gesture recognition such as template matching (Hernandez-Rebollar, Kyriakopoulos, & Lindeman, 2004; Keni, Meher, & Marathe, 2013) and some used machine learning classifiers such as Hidden Markov Model (HMM) which would be further discussed in the next section. According to Khan and Ibraheem (2012), Hand Gesture Recognition could be applied in different applications, including sign language recognition system, robot control, graphic editor, virtual environments (VE), numbers recognition, television control, and 3D modelling. These applications implemented hand gesture recognition system methods to perform the basic function of the application. One example of a sign language recognition system utilized hand gesture recognition module. After the image/video is recorded, the stream of data is loaded and segmented. Then, feature extraction is conducted in regards of shape, size, hand and finger trajectory, and body of the signer. HMM requires feature vectors; therefore, feature vector is created by using this information. Then, classification is performed by Hidden Markov Model. Feature vector must be extracted from each video frame for both training and recognition, then inputted into the Hidden Markov Model (Yang, Tao, Xi, & Ye, 2015).

METHODS

Classification Methods and Related Works

Hidden Markov Model (HMM). Several previous studies implement Hidden Markov Model (HMM) as the basis of research to make Sign Language Recognition (SLR) system. HMM is a statistic model where a set of parameters is hidden. The hidden parameters can be acquired from related observation parameters (Zabulis, Baltzakis, & Argyros, 2009). "Hidden Markov Model is a finite model that describes a probability distribution over an infinite number of sequences" (Eddy, 1996). HMM has been widely employed in Speech Recognition system. HMM is also used in glove-based Sign Language Recognition system (Liang & Ouhyoung, 1998; Ma, Gao, Wu, & Wang, 2000; H. Wang, Leu, & Oz, 2006). Hidden Markov Model is used to tackle sequential data and Sign Language consists of continuous gestures which make up a word or sentence.

Multi-dimensional Hidden Markov Model is used in recognizing American Sign Language (ASL) in, which has 96.7% accuracy (Wang et al., 2006). The data from input devices (CyberGlove™) is in the form of 21 data-stream, which is then segmented into gestures in the same interval. Subsequently, the data is inputted into 21-dimensional feature vector. Finally, the data is classified and recognized according to stochastic data to produce the output (recognized sign language). Similar previous research (Ma et al., 2000) also used HMM to recognize Chinese Sign Language (CSL) and was able to get 98.2% accuracy with embedded training. In an earlier research, Liang and Ouhyoung (1998) used HMM to recognize data from DataGlove™. The system could recognize 250 words of Taiwanese Sign Language (TSL) formed by 51 basic postures, 6 orientations, and 8 fundamental motions with 80.4% accuracy of real-time continuous gestures. The input was statistically analyzed by 4 parameters. The parameters are position, posture, motion, and orientation.

Starner and Pentland (1997) also created a real-time SLR system. However, their research used video-based approach. The research conducted two experiments by using 40 words lexicon, where in the first experiment the user wore coloured-glove and in the second experiment without coloured-glove. The first experiment reached 99% of accuracy, whilst the second experiment reached 92% of accuracy. Elmezain, Al-Hamadi, Appenrodt, & Michaelis (2008) created a real-time SLR system for 10 Arabic numbers in Sign Language using HMM from coloured image sequence and attained 98.94% of accuracy. There are SLR systems using HMM from various countries. The previous research includes Sign Language Recognition system for Taiwanese Sign Language (Lee, Yeh, & Hsiao, 2016; Liang & Ouhyoung, 1998), Chinese Sign Language (Ma et al., 2000), American Sign Language (Starner & Pentland, 1997; Wang et al., 2006), and Arabic Sign Language (Elmezain et al., 2008). Moreover, there is also Greek Sign Alphabet Letters Recognition system (Pashaloudi & Margaritis, 2004), which is able to recognize 90.20% of training set and 86.52% of testing set. The sets include 16 Greek Sign Alphabet Letters and the system recognizes pictures. Hidden Markov Model struggles handling noisy data (Roh et al., 2016); therefore, Kaluri & Pradeep, (2017) proposed the implementation of Wiener Filter to eliminate the noise in an images and Adaptive Histogram technique to segment the images which would be feed into HMM for training and recognition.

Modifications of Hidden Markov Model. Modification of Hidden Markov Model (HMM) has been researched to better improve the performance and accuracy of a Sign Language Recognition (SLR) system. For instance, a 3 state left-to-right Hidden Markov Model with three independent Gaussian Mixtures (GMM) and a globally merged covariance matrix is implemented in research with 17% error rate (Dreuw, Rybach, Deselaers, Zahedi, & Ney, 2007). This research implements speech recognition techniques to create an automatic Sign

Language Recognition system which can adapt to dialects in the sign language. Principal Component Analysis (PCA) is used to lessen the dimension and help the classification process of features obtained. "Principal Component Analysis is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set" (Ringnér, 2008). PCA is commonly used in hand recognition systems (Li, Kao, & Kuo, 2016; Prasad, Kishore, Kumar, & Kumar, 2016; Sawant, 2014; Zaki & Shaheen, 2011). PCA is also used to characterize the feature of fingers in a vision-based Chinese SLR system (Zhang et al., 2004). Tied-Mixture Density Hidden Markov Model (TMDHMM) is used to speed up the recognition system without significantly reducing the accuracy. The system could recognize up to 92.5% of the frequently used Chinese Sign Language. TMDHMM is also used because of the efficient computational costs. In (Yang et al., 2015), Weighted Hidden Markov Model assigns weights for each sign samples. This system used Kinect to improve recognition rate with 156 isolated sign words used as data. It attained a high recognition rate of up to 94.74%.

The latest studies have several Hidden Markov Model based improvements on the previous research. The research includes a hybrid CNN-HMM (Koller, Zargaran, Ney, & Bowden, 2016), Coupled-HMM (CHMM) (Kumar, Gauba, Roy, & Dogra, 2017), and Kinect-Based using Hidden Markov Model (Lee et al., 2016) for SLR system. CHMM was used in an SLR system and was able to reach 90.80% of accuracy (Kumar et al., 2017). This research used Kinect to create 3D model of the captured gestures. Another research also used Kinect for HMM based Taiwanese Sign Language Recognition (SLR) system and was able to attain 85.14% of recognition rate (Lee et al., 2016). HMM was used to determine the signing direction, whilst for the recognition, a trained SVM was used.

Table 1 summarizes the SLR using HMM and modified HMM. As shown in the table, the highest reported accuracy was achieved by Elmezain et al. (2008) on a small size dataset. The system could recognize 0-9 Arabic numbers in Arabic Sign Language. Some of the system could only recognize alphabets, numbers, and basic hand shapes (Elmezain et al., 2008; Pashaloudi & Margaritis, 2004; Wang et al., 2006) with high accuracy. There are also Sign Language Recognition system which can recognize a large number of vocabularies (Dreuw et al., 2007; Liang & Ouhyoung, 1998; Ma et al., 2000; Zhang et al., 2004). The system could recognize 220 words and 80 sentences in Chinese Sign Language with 98.2% of accuracy (Ma et al., 2000). However, one of the best systems from all of the systems mentioned before could recognize 92.5% of 439 words from Chinese Sign Language (Zhang et al., 2004). The sample used was 1756 words for training and 439 for testing. Weighted Hidden Markov Model is also worth mentioning as it outperformed other methods as well as the traditional Hidden Markov Model (Yang et al., 2015).

Table 1
Summary of SLR using HMM and modified HMMs

Author	Method(s)	Data	Result (%)
(Wang et al., 2006)	Multi-dimensional HMM	26 ASL alphabets and 36 ASL handshapes	96.7
(Ma et al., 2000)	HMM	220 CSL words and 80 CSL sentences	98.2
(Liang & Ouhyoung, 1998)	HMM	250 TSL words	80.4
(Starner & Pentland, 1997)	HMM	40 ASL words lexicon	92 – 99
(Elmezain et al., 2008)	**HMM**	10 numbers in Arabic Sign Language	**98.94**
(Pashaloudi & Margaritis, 2004)	HMM	16 letters in GSL	86.52 - 90.20
(Dreuw et al., 2007)	3 states left-to-right HMM	RWTH-Boston-104 corpus (201 sequence, 104 words)	17 (error rate)
(Zhang et al., 2004)	TMDHMM	439 signs in CSL	92.5
(Koller et al., 2016)	CNN-HMM	-	-
(Kumar et al., 2017)	CHMM	25 dynamic words from Indian Sign Language	90.80
(Lee et al., 2016)	HMM-SVM	12 directions from TSL	85.14
(Yang et al., 2015)	WHMM	156 isolated signs in CSL	94.74

Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is parallel computational models that simulate human brain, where processing nodes are called "neurons". Every neurons stored information and depend on it, it receives input from and send output to other neurons (Dongare, Kharde, & Kachare, 2012). (Adithya, Vinod, & Gopalakrishnan, (2013) employed ANN forward-backward algorithm to automatically recognize 26 alphabets and 10 numbers of Indian Sign Language with 91.1% of accuracy. However, the system only worked statically, not in real-time situation. Backpropagation Neural Network is the family of neural network models, where the learning algorithm is based on Deepest-Descent technique (Buscema, 1998). It was employed to build Sign Language Recognition (SLR) system for Indian Sign Language and attained 92.34% of recognition rate (Prasad, Kishore, & Kumar, 2016). The data used are own-made 80 sequences of video which in total consist of 59 signs of letters, numbers, and 23 words in Indian sign language.

Convolutional Neural Network (CNN)

"Convolutional Neural Network is the family of neural network models that feature a type of layer known as the convolutional layer which can extract features" (Pham, Kruszewski, & Boleda, 2016). CNN has been widely used in Computer Vision projects. Several of those projects used CNN to recognize Sign Language. Huang et al. (2015) used 3D CNN to

recognize Sign Language into texts or speech. The accuracy of the system was as high as 94.2%. The data used consisted of 25 words of sign language used in daily conversations. Pigou, Dieleman, Kindermans, & Schrauwen, (2014) used CNN for Italian SLR. The system was automated due to Convolutional Neural Network and used Microsoft Kinect and GPU acceleration. The system could recognize 20 Italian Gestures with the accuracy of 91.7%. A Hybrid CNN-Hidden Markov Model (CNN-HMM) was used for continuous SLR system (Koller et al., 2016). Hidden Markov Model has the capabilities of sequence modelling. It can be combined with CNN, which has discriminative strength. The best system using CNN had 94.2% of recognition rate; however, the data used was small (Huang et al., 2015). 3D Convolutional Neural Network is employed in the creation of LipNet, which is a deep neural network that could do lipreading through visual approach. For the first time, LipNet could recognize end-to-end sentence-level lipreading and achieved state-of-the-art result of 95.5%. LipNet architecture consisted of 3D convolutional Neural Network, Bidirectional RNN, SOFTMAX, and Connectionist Temporal Classification (CTC). The experiments were performed on GRID corpus in which 28775 videos were used for training and 3971 videos were used for testing (Assael, Shillingford, Whiteson, & de Freitas, 2016).

Self-organizing Map (SOM)

"The Self-Organizing Map (SOM) is a software tool for the visualization of high-dimensional data. It implements an orderly mapping of a high-dimensional distribution onto a regular low-dimensional grid" as defined by (Kohonen, 1998). A Kohonen Self-Organizing Map was used in an Indian Sign Language Recognition system. SOM is sometimes also called Kohonen Self-Organizing Feature Map (SOFM). Neural Network is used for pattern recognition run in MATLAB. The system attained a maximum accuracy of 80% from 35 images of 5 sign language captured in 7 different backgrounds, with 72 seconds of training time. Gao et al. (2004) employed Self-Organizing Map together with other methods such as Hidden Markov Model (HMM) and Simple Recurrent Network (SRN). Simple Recurrent Network was implemented to segment continuous sign language according to Self-Organizing Feature Map (SOFM) representations. The output of Simple Recurrent Network (SRN) was then used as HMM states. Viterbi Algorithm is employed to search the best matched word of the according signs. The highest accuracy was 91.3% in embedded training and registered test sets.

Other Methods

Korean Sign Language Recognition system was able to attain 85% of recognition rate out of 25 Korean Sign Language (KSL) word (Kim, Jang, & Bien, 1996). The system employed Fuzzy Min-Max to recognize the input from a Data-glove. The data received was the position of hand (x, y, z axes) forming 10 primary movements and 14 basic hand

shapes. “Fuzzy Min-Max Classification Neural Network is Hyper boxes defined by pairs of min-max points, and their corresponding membership functions are used to create fuzzy subsets of the n-dimensional pattern space” as defined by (Simpson, 1992). Template Matching is often used in Computer Vision, also used in Sign Language Recognition system (Hernandez-Rebollar et al., 2004; Keni et al., 2013). Helped by 17 volunteers with different level of skills to demonstrate 30 American Sign Language words, the system in (Hernandez-Rebollar et al., (2004) attained 98% of accuracy. Ghosh and Ari (2016) proposed an enhanced version of Radial Basis Function (RBF) Neural Network for the classification of hand gestures. K-means algorithm was used to select the centers of RBF classifier automatically and least-mean-square (LSM) algorithm is utilized to update the estimated weight matrix recursively. K-Nearest Neighbors (KNN) classifier was used along with PCA, generating 96.31% highest accuracy. The system consisted of three phases: hand segmentation, feature extraction, and classification. To capture the image depth, the system used Kinect. The database used consisted of 61 hand gestures, each from 10 different signers totaling 12,200 images of Brazilian Sign Language (Costa Filho, Souza, Santos, Santos, & Costa, 2017). KNN classifier also produces a high result of 99.61% accuracy in recognizing Indian Sign Language. The data used consisted of 3600 images from 40 different signers. The data varies in lighting condition, angle, and distance. Lim, Tan and Tan (2016) handled the problems in sign language recognition using feature covariance matrix to recognize isolated sign language with 87.33% recognition rate for ASL. Feature covariance matrix is able to reduce the dimension of features and combine associated signs naturally. Roh et al. (2016) proposed a novel approach in handling noisy data by using Selective Temporal Filtering (STF). In a not noisy environment, the system reached 92.1% of accuracy; meanwhile, in a noisy environment, the accuracy dropped to 62.2%. Their study found that STF performs better than HMM and Conditional Random Fields not only in a noisy environment but also in a balanced environment.

RESULTS AND DISCUSSION

There are several notable prior works that represent the direction of the classification methods used in SLR system. Ma et al. (2000) proposed a HMM based sign language recognizer that could achieve high accuracy results in recognizing Chinese Sign Language sentences. Gesture inputs were obtained using two Cybergloves and two 3SAPCE-position trackers, and invariant features were extracted based on 3D movements of the signer. Using the combination of dynamic programing, HMM, Bigram language model, and search algorithm, they managed to recognize CSL at a sentence level at 98.2% accuracy. However, their SLR system still relied on active sensors for data acquisition to achieve high accuracy results. HMM-based SLR approaches have been shown to achieving good recognition accuracy especially in small to medium-sized datasets. Efforts on improving

the performance of HMM-based approaches have also been proposed by modifying the standard HMM method. Zhang et al. proposed the tied-mixture density HMM (TMDHMM) to increase the recognition speed without sacrificing the recognition accuracy (Zhang et al., 2004). Meanwhile, Yang et al. (2015) proposed weighted HMM (WHMM) to cope with sign variations from different signers. Both modifications can achieve more than 90% accuracy on medium-sized datasets of word-level signs. Nevertheless, the proposed HMM-based and modified HMM-based approaches are still required to transform the input data into "handcrafted" sign features before applying the classification method. Designing invariant sign features can be tedious works and highly dependent on the type of input data being used. Moreover, feature extraction can also be seen as an additional step potentially slowing down the recognition speed and the classification performance often depends on the quality of the extracted sign features.

Koller et al. (2016) proposed a Hybrid CNN-HMM approach in order to combine the discriminative strength of the CNN and the sequential modeling of the HMM. In this hybrid approach, they used the output posteriors of the CNN as observation probabilities of the HMM. Therefore, it enabled them to perform end-to-end training. They also compared the performance of the Hybrid CNN-HMM to the Tandem CNN-HMM approach, where CNN was treated as a feature extractor for the HMM classifier. Their experimental results showed that the Hybrid CNN-HMM outperformed the Tandem CNN-HMM approach. The current implementation of the hybrid approach still focuses only on the right hand, which is assumed to be the dominant hand of the signer. This method also requires a good frame-state alignment. An attempt to break away from HMM-based approach was done by Huang et al. (2015). They proposed a 3D CNN to recognize sign language directly from input data without the need of designing handcrafted features. The input data was obtained from a Microsoft Kinect, which contained color and depth information, and the body joint positions of the signer. The proposed method had shown to achieve promising results. However, the results were obtained from a small-sized dataset. This approach also cannot handle variable-length sequences of frames. Additional pre-processing is necessary to select a fixed number of frames for each input channel. Recently, a promising deep learning approach for sentence-level sequence prediction has been proposed by Assael, Shillingford, Whiteson and de Freitas (2016). They combined the 3D CNN and Bidirectional RNN to perform lipreading at the sentence-level. The proposed approach was evaluated on large-sized dataset and achieved 95.5% accuracy. Lipreading task is very similar to sign language recognition where both try to interpret spatio-temporal cues obtained from input video to predict spoken words or sentences. A similar approach can be employed to solve SLR problem. However, it tends to overfit when trained on small dataset due to complexity of the deep model. Also, it requires high computational load. It is of great interest to develop a simpler network that can achieve high performance and requires low

computational load, but still embeds the feature learner into the classifier in multi-layered neural network fashion.

CONCLUSION AND FUTURE WORK

Different approaches have been proposed by various studies to solve the problem of sign language recognition (SLR). We have reviewed prior works based on different stages of the recognition procedure, which includes data acquisition, gesture segmentation, feature extraction, and classification. In this contribution, we provide more coverage on the classification approaches. We have reviewed many prior works such as HMM-based, modified HMM based, neural network based, and hybrid-based approaches. HMM-based SLR approaches have been shown to achieving good recognition accuracy especially in small to medium sized datasets. Efforts on improving the performance of HMM-based approaches have also been proposed by modifying the standard HMM method. However, the proposed HMM-based and modified HMM-based approaches still require the extraction of sign features from input data before applying the classification method. Designing invariant sign features can be tedious works highly dependent on the type of input data being used. Moreover, feature extraction also contributes to the computational load and the classification performance often depends on the quality of the extracted sign features.

Combinations of CNN and HMM have also been proposed to improve the performance of the SLR system. It has been shown that a hybrid approach that embeds CNN into HMM abiding to Bayesian principles outperforms the tandem approach that treats CNN as a feature extractor. The hybrid approach also enables the end-to-end training. The 3D CNN and the combination of 3D CNN and RNN have recently been shown to potentially improve the SLR performance. However, overfitting and high computational requirements still hinder their adoption. We believe the future direction of the research is toward developing a simpler network that can achieve high performance and requires low computational load, which embeds the feature learner into the classifier in multi-layered neural network fashion.

ACKNOWLEDGEMENT

This research was supported by Research Grant No. 039A/VR.RTT/VI/2017 from Ministry of Research Technology and Higher Education of Republic of Indonesia.

REFERENCES

Adithya, V., Vinod, P. R., & Gopalakrishnan, U. (2013). Artificial neural network based method for Indian sign language recognition. In *2013 IEEE Conference on Information & Communication Technologies (ICT)* (pp. 1080–1085). Thuckalay, Tamil Nadu, India.

Ahuja, M. K., & Singh, A. (2015). A Survey of Hand Gesture Recognition. *International Journal*, *3*(5), 266-271.

Apoorva Ranjini, S. S., Chaitra, M., Deepika, V., & Patil, J. M. (2014). Sign Language Recognition System. *International Journal on Recent and Innovation Trends in Computing and Communication*, *2*(4), 947 – 953.

Assael, Y. M., Shillingford, B., Whiteson, S., & de Freitas, N. (2016). LipNet: Sentence-level Lipreading. Computing Research Repository (CoRR), 1611.01599, 1-13. Retrieved July 4, 2017 from http://arxiv.org/abs/1611.01599.

Bhavsar, H. (2017). Review on Classification Methods used in Image based Sign Language Recognition System. *International Journal on Recent and Innovation Trends in Computing and Communication*, *5*(5), 949–959.

Bhavsar, H., & Trivedi, J. (2017). Review on Feature Extraction methods of Image based Sign Language Recognition system. *Indian Journal of Computer Science and Engineering (IJCSE), 8*(3), 249-259.

Bhuyan, M. K., Kumar, D. A., MacDorman, K. F., & Iwahori, Y. (2014). A novel set of features for continuous hand gesture recognition. *Journal on Multimodal User Interfaces*, *8*(4), 333–343.

Buscema, M. (1998). Back propagation neural networks. *Substance Use & Misuse*, *33*(2), 233–270.

Carneiro, S. B., Ferreira, J. O., Barbosa, T. M., Da Rocha, A. F., Soares Alcala, S. G., & M. Santos, E. D. (2016). Static Gestures Recognition for Brazilian Sign Language with Kinect Sensor. In *2016 IEEE SENSORS* (pp. 1-3). Orlando, FL, USA.

Chai, X., Li, G., Lin, Y., Xu, Z., Tang, Y., Chen, X., & Zhou, M. (2013). Sign language recognition and translation with kinect. In *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 1-2). Shanghai, China.

Costa Filho, C. F. F., Souza, R. S. de, Santos, J. R. dos, Santos, B. L. dos, & Costa, M. G. F. (2017). A fully automatic method for recognizing hand configurations of Brazilian sign language. *Research on Biomedical Engineering*, *33*(1), 78–89.

Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology*, *2*(1), 189–194.

Dreuw, P., Rybach, D., Deselaers, T., Zahedi, M., & Ney, H. (2007). Speech recognition techniques for a sign language recognition system. In *Eighth Annual Conference of the International Speech Communication Association* (pp. 2513-2516). Antwerp, Belgium.

Eddy, S. R. (1996). Hidden markov models. *Current Opinion in Structural Biology*, *6*(3), 361–365.

Elmezain, M., Al-Hamadi, A., Appenrodt, J., & Michaelis, B. (2008). A hidden markov model-based continuous gesture recognition system for hand motion trajectory. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on* (pp. 1–4). Tampa, FL, USA.

Escobedo, E., & Camara, G. (2016). A new Approach for Dynamic Gesture Recognition using Skeleton Trajectory Representation and Histograms of Cumulative Magnitudes. In *SIBGRAPI Conference on Graphics, Patterns and Images* (pp. 209–216). São Paulo, Brazil.

Eugene Starner, T. (1995). *Visual Recognition of American Sign Language Using Hidden Markov Models*. Massachusetts Inst Of Tech Cambridge Dept Of Brain And Cognitive Science.

Gao, W., Fang, G., Zhao, D., & Chen, Y. (2004). A Chinese sign language recognition system based on SOFM/SRN/HMM. *Pattern Recognition*, *37*(12), 2389–2402.

Ghosh, D. K., & Ari, S. (2016). On an algorithm for Vision-based hand gesture recognition. *Signal, Image and Video Processing*, *10*(4), 655–662.

Goyal, E. K., & Singh, A. (2014). Indian Sign Language Recognition System for Deaf People. *Journal on Today's Ideas - Tomorrow's Technologies*, *2*(2), 145–151.

Goyal, S., Sharma, I., & Sharma, S. (2013). Sign Language Recognition System for Deaf and Dumb People. *International Journal of Engineering Research and Technology*, *2*(4), 382–387.

Gurjal, P., & Kunnur, K. (2012). Real time hand gesture recognition using SIFT. *International Journal of Electronics and Electrical Engineering*, *2*(3), 19–33.

Hernandez-Rebollar, J. L., Kyriakopoulos, N., & Lindeman, R. W. (2004). A new instrumented approach for translating American Sign Language into sound and text. In *Sixth IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 547–552). Seoul, South Korea.

Huang, J., Zhou, W., Li, H., & Li, W. (2015). Sign language recognition using 3D convolutional neural networks. In *Multimedia and Expo (ICME), 2015 IEEE International Conference on* (pp. 1–6). Turin, Italy.

Jiang, Y., Tao, J., Ye, W., Wang, W., & Ye, Z. (2014). An Isolated Sign Language Recognition System Using RGB-D Sensor with Sparse Coding. In *2014 IEEE 17th International Conference on Computational Science and Engineering* (pp. 21–26). Chengdu, China.

Kakde, M. U., Nakrani, M. G., & Rawate, A. M. (2016). A Review Paper on Sign Language Recognition System For Deaf And Dumb People using Image Processing. *International Journal of Engineering Research and Technology*, *5*(3), 590–592.

Kalsh, E. A., & Garewal, N. S. (2013). Sign Language Recognition System. *International Journal of Computational Engineering Research*, *3*(6), 15-21.

Kaluri, R., & Pradeep, C. H. (2017). An enhanced framework for sign gesture recognition using hidden Markov model and adaptive histogram technique. *International Journal of Intelligent Engineering & Systems, 10*(3), 11-19.

Keni, M., Meher, S., & Marathe, A. (2013). Sign Language Recognition System. *International Journal of Scientific and Engineering Research*, *4*(12), 580-583.

Keskin, C., Kırac, F., Kara, Y. E., & Akarun, L. (2011) Real Time Hand Pose Estimation using Depth not according to formarSensors. In *IEEE International Conference on Computer Vision Workshops* (pp. 1228–1234). Barcelona, Spain.

Khan, R. Z., & Ibraheem, N. A. (2012). Hand gesture recognition: a literature review. *International Journal of Artificial Intelligence & Applications*, *3*(4), 161.

Kim, J. S., Jang, W., & Bien, Z. (1996). A Dynamic Gesture Recognition System For The Korean Sign Language (KSL). *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *26*(2), 354–359.

Kishore, P. V. V, Prasad, M. V. D., Prasad, C. R., & Rahul, R. (2015). 4-Camera model for sign language recognition using elliptical fourier descriptors and ANN. In *International Conference on Signal Processing and Communication Engineering Systems - Proceedings of SPACES 2015, in Association with IEEE* (pp. 34–38). Guntur, India.

Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, *21*(1), 1–6.

Koller, O., Zargaran, O., Ney, H., & Bowden, R. (2016). Deep Sign: Hybrid CNN-HMM for Continuous Sign Language Recognition. In *Proceedings of the British Machine Vision Conference 2016*. New York, UK.

Koul, M., Patil, P., Nandurkar, V., & Patil, S. (2016). Sign Language Recognition Using Leap Motion Sensor. *International Research Journal of Engineering and Technology (IRJET)*, *3*(11), 322–325.

Kumar, P., Gauba, H., Roy, P. P., & Dogra, D. P. (2017). Coupled hmm-based multi-sensor data fusion for sign language recognition. *Pattern Recognition Letters*, *86*, 1–8.

Lee, G. C., Yeh, F.-H., & Hsiao, Y.-H. (2016). Kinect-based taiwanese sign-language recognition system. *Multimedia Tools and Applications*, *75*(1), 261–279.

Li, T. H. S., Kao, M. C., & Kuo, P. H. (2016). Recognition System for Home-Service-Related Sign Language Using Entropy-Based K-Means Algorithm and ABC-Based HMM. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *46*(1), 150–162.

Liang, R. H., & Ouhyoung, M. (1998). A real-time continuous gesture recognition system for sign language. In *Third IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 558–567). Nara, Japan.

Lim, K. M., Tan, A. W. C., & Tan, S. C. (2016). A feature covariance matrix with serial particle filter for isolated sign language recognition. *Expert Systems with Applications*, *54*, 208–218.

Ma, J., Gao, W., Wu, J., & Wang, C. (2000). A Continuous Chinese Sign Language Recognition System. In *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000* (pp. 428–433). Washington, DC, USA

Majid, M. A., & Zain, J. M. (2013). A Review on The Development of Indonesian Sign Language Recognition System. *Journal of Computer Science*, *9*(11), 1496–1505.

Mehdi Y. N., S. A. K. (2002). Sign language recognition using sensor gloves. In *Proceedings of the 9th International Conference on Neural Information* (pp. 2204–2206). Singapore.

Mohandes, M., Aliyu, S., & Deriche, M. (2014). Arabic sign language recognition using the leap motion controller. In *23rd International Symposium on Industrial Electronics (ISIE)* (pp. 960–965). Istanbul, Turkey.

Ong, S. C. W., & Ranganath, S. (2005). Automatic Sign Language Analysis : A Survey and the Future beyond Lexical Meaning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(6), 873–891.

Pankajakshan, P. C., & Thilagavathi, B. (2015). Sign Language Recognition System. In *International conference on Innovation in Information Embedded and Communication System* (pp. 2–5). Coimbatore, India.

Pashaloudi, V. N., & Margaritis, K. G. (2004). A performance study of a recognition system for greek sign language alphabet letters. In *9th Conference Speech and Computer* (pp. 545-551). Saint-Petersburg, Russia.

Pham, N. Q., Kruszewski, G., & Boleda, G. (2016). Convolutional Neural Network Language Models. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 1153–1162). Austin, Texas.

Phi, L. T., Nguyen, H. D., Bui, T. T. Q., & Vu, T. T. (2015). A glove-based gesture recognition system for Vietnamese sign language. *Automation and Systems, Proceedings*, *13*(16), 1555–1559.

Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2014). Sign language recognition using convolutional neural networks. In *Workshop at the European Conference on Computer Vision* (pp. 572–578). Cham : Springer.

Pisharady, P. K., & Saerbeck, M. (2015). Recent methods and databases in vision-based hand gesture recognition: A review. *Computer Vision and Image Understanding*, *141*, 152–165.

Potter, L. E., Araullo, J., & Carter, L. (2013). The Leap Motion controller. In *Proceedings of the 25th Australian Computer-Human Interaction Conference on Augmentation, Application, Innovation, Collaboration* (pp. 175–178). Adelaide, Australia.

Prasad, D., Kishore, V., & Kumar, A. (2016). Indian sign language recognition system using new fusion based edge operator. *Journal of Theoretical and Applied Information Technology*, *88*(3), 574–584.

Raheja, J. L., Mishra, A., & Chaudhary, A. (2016). Indian Sign Language Recognition Using SVM 1. *Pattern Recognition and Image Analysis*, *26*(2), 434–441.

Rakun, E., Andriani, M., Wiprayoga, I. W., Danniswara, K., & Tjandra, A. (2013). Combining depth image and skeleton data from kinect for recognizing words in the sign system for indonesian language (sibi [sistem isyarat bahasa indonesia]). In *2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS)* (pp. 387–392). Bali, Indonesia.

Ringnér, M. (2008). What is principal component analysis? *Nature Biotechnology*, *26*(3), 303.

Roh, M. C., Fazli, S., & Lee, S. W. (2016). Selective temporal filtering and its application to hand gesture recognition. *Applied Intelligence*, *45*(2), 255–264.

Saengsri, S., Niennattrakul, V., & Ratanamahatana, C. A. (2012). TFRS: Thai finger-spelling sign language recognition system. In *2nd International Conference on Digital Information and Communication Technology and Its Applications, DICTAP* (pp. 457–462). Bangkok, Thailand.

Sahoo, A. K., Mishra, G. S., & Ravulakollu, K. K. (2014). Sign language recognition: state of the art. *ARPN Journal of Engineering and Applied Sciences*, *9*(2), 116–134.

Sawant, S. N. (2014). Sign Language Recognition System to aid Deaf-dumb People Using PCA. *International Journal of Computer Science and Engineering Technology*, *5*(5), 570–574.

Simpson, P. K. (1992). Fuzzy min-max neural networks. I. Classification. *IEEE Transactions on Neural Networks*, *3*(5), 776–786.

Starner, T., & Pentland, A. (1997). Real-time american sign language recognition from video using hidden markov models. In *Motion-Based Recognition* (pp. 227–243). Dordrecht: Springer.

Wang, H., Leu, M. C., & Oz, C. (2006). American Sign Language Recognition Using Multi-dimensional Hidden Markov Models. *Journal of Information Science and Engineering*, *22*(5), 1109–1123.

Wang, R. Y., & Popović, J. (2009). Real-time hand-tracking with a color glove. In *ACM transactions on graphics (TOG)* (Vol. 28, p. 63). New Orleans, Louisiana.

Yan, S., Xia, Y., Smith, J. S., Lu, W., & Zhang, B. (2017). Multiscale Convolutional Neural Networks for Hand Detection. *Applied Computational Intelligence and Soft Computing*, *2017*, 1-13.

Yang, W., Tao, J., Xi, C., & Ye, Z. (2015). Sign Language Recognition System Based on Weighted Hidden Markov Model. In *8th International Symposium on Computational Intelligence and Design (ISCID)* (pp. 449–452). Hangzhou, China.

Zabulis, X., Baltzakis, H., & Argyros, A. (2009). Vision-based hand gesture recognition for human-computer interaction. In *The universal access handbook* (pp. 1–30). Boca Raton: CRC Press.

Zaki, M. M., & Shaheen, S. I. (2011). Sign language recognition using a combination of new vision based features. *Pattern Recognition Letters*, *32*(4), 572–577.

Zhang, L. G., Chen, Y., Fang, G., Chen, X., & Gao, W. (2004). A Vision-based Sign Language Recognition System Using Tied-mixture Density HMM. In *Proceedings of the 6th international conference on Multimodal interfaces* (pp. 198–204). State College, Pennsylvania, USA.

Pertanika J. Sci. & Technol. 26 (4): 1677 - 1692 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Physical Layer Security with RF Energy Harvesting Protocols for Wireless Networks

Jindal, P.* and Sinha, R.

Department of Electronics and Communication Engineering, National Institute of Technology, Kurukshetra-136119, India

ABSTRACT

In this paper, the secrecy performance of two energy harvesting (EH) protocols, known as, time switching (TS) and power splitting (PS) is analyzed for a single hop relaying network, in which the sender sends the data to legitimate destination using a relay in the presence of an eavesdropper. In the considered network, the relay and source are powered up by using energy harvesting techniques. The performance of the systems with EH is compared with that of the conventional physical layer security model, where nodes are powered up by individual battery sources. The secrecy rates with the two different types of EH protocols and conventional system are analyzed for two relaying schemes: decode-and-forward (DF), and amplify-and-forward (AF). Resulting analysis shows that the TS EH system has higher secrecy rate as compared to conventional system and the secrecy rate of the conventional system is higher than that of PS EH protocol.Further, the simulation results show that AF relays outperforms DF relays in all the scenarios.

Keywords: Amplify and forward, energy harvesting, jamming, physical layer security, secrecy rate

INTRODUCTION

Physical layer security (PLS) and energy harvesting have been gaining a phenomenal growth and attention from the research community. PLS guarantees and enhance confidentiality of the transmitted information (Wyner, 1975), whereas EH utilizes the harvested energy for information processing (Hoang, Duong, Vo, & Kundu, 2017). Although an in-depth investigation of both technologies as an individual body of knowledge has been done in the literature, but the combination of both has been recently recurred as an attractive research paradigm. Physical layer

ARTICLE INFO

Article history:
Received: 30 August 2017
Accepted: 27 June 2018
Published: 24 October 2018

E-mail addresses:
poonamjindal81@nitkkr.ac.in (Jindal, P.)
rupalisinha468@gmail.com (Sinha, R.)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

security helps in providing secure data transmission by using physical characteristics of the wireless medium.

Various techniques are used to support PLS such as jamming, full duplex operations, interference cancellation, relaying schemes such as DF and AF (Dong, Han, Petropulu, & Poor, 2010). Conventionally, nodes in the wireless networks perform their respective roles using individual battery sources. However, in some situations, it is not possible to recharge or replace these batteries (Zhan, & Ho, 2013).

Sometimes, especially in wireless sensor networks, it is hazardous to replace these batteries. Moreover, it is the need of the hour to adopt energy efficient systems as a concern towards environment and conserve energy as much as much as possible as disposal of batteries is also not environment friendly. Different sectors of information and communication technologies (ICT) greatly contribute in CO_2 emission and huge energy consumption. In 2012, about 4.7% of world's electrical energy was consumed by ICT (approx. 920 TWh, 1 TWh=10^{12} Watt hour) and releasing around 1.7% of total carbon emission into the atmosphere. Saving energy to prolong the lifetime of energy constrained nodes, maintaining the parameter metric i.e. secrecy rate, throughput, outage probability, optimize the power allocation to nodes, maintaining the confidentiality, authentication during transmission between legitimate nodes are some of the critical issues in wireless networks.

EH techniques helps in solving energy issues to some extent (Yuen, Elkashlan, Qian, Duong, Shu, & Schmidt, 2015 Part I-III). The two commonly used EH protocols are PS and TS (Nasir, Zhou, Durrani, & Kennedy, 2013). Recently, EH technique has gained considerable attention, especially in energy-constrained networks. This technique involves conversion of the received radio frequency (RF) signal into electricity. It helps wireless devices so that it can harvest energy from RF signals for processing and transmission of information simultaneously. Concept of simultaneous wireless information power transfer (SWIPT) has attracted researchers to replenish energy limited devices. Radio frequency (RF) signal can carry information and energy simultaneously and thus motivating green communication for energy scavenging. Hence, green communication has emerged as latest technology to power next generation wireless networks.

PLS provides the confidentiality to the information. When it combines with the cooperative schemes, enhances the signal coverage along with secrecy. Further, when both of the techniques are combined with RF EH technology, enhances the lifetime of battery-constrained devices. The combination of three techniques; providing information security while maintaining the energy of wireless nodes is considered in this paper. Therefore, the scenario of PLS with energy deprived relay nodes is taken into consideration. The beacon node allocates energy to source and relay node. TS and PS protocols are used to evaluate the secrecy rate of system. In TS, time switching factor is decided to perform energy

harvesting (EH) and information processing (IP). In PS, some portion of received power is used for EH, while remaining power for IP by the relay. Further, EH techniques have been implemented using AF and DF relaying techniques. The performance of systems employing EH techniques is compared with that of the conventional system in terms of secrecy rate, the rate at which information can be sent securely from sender to the intended receiver. An in-depth analysis of different energy harvesting protocols with physical layer security in different cooperative scenarios makes this paper different from the existing work, as the existing literature does not present such a deep insight and an extensive comparative analysis of the problem under study.

Related Work

Some recent studies have investigated the protocols used in EH for transferring confidential information through PLS. For secure communication, power splitting receivers are used for evaluating performance of DF and AF protocols using secrecy outage probability (Son, & Kong, 2015). EH based AF multi antenna, HD relay network with EH relaying protocols: TSR, PSR and IRR in presence of passive eavesdropper is studied. The expression for ergodic secrecy capacity is derived and factors affecting secrecy capacity are investigated. Time switching ratio and power splitting factor for TS and PS are further calculated. (Salem, A., Hamdi, & Rabie, 2016). Two-phase communication protocol for wireless powered cooperative jammer is proposed where energy is harvested by signal transmitted by source. High SNR and antenna regime are considered for increasing throughput with fixed rate transmission.Throughput maximization is done by optimal time allocation for PS and TS protocols. (Liu, Zhou, Durrani, & Popovski, 2016). Wireless power transfer schemes for source, destination and joint destination and source are analyzed. Further, two suboptimal strategies for WPT are proposed to maximize information throughput (Zhang, & Chen, 2017). FD information source capable of self energy recycling, powered by energy source for transmitting confidential information is given (Wu, Wang, Deng, & Zhang, 2017).

Need of Cooperative Physical Layer Security with RF Energy Harvesting

The area under study has a number of real time applications. The upcoming wireless networks including IoTs, 5G networks are based on sensor nodes. Extensive amount of sensitive information is collected and transmitted from wireless sensor nodes that are powered up by individual battery source and their energy is depleted with the passage of time. The proposed schemes are beneficial in extending the battery lifetime while maintaining the information security. The work presented in this paper is beneficial in the military applications where, the soldiers work at hard to reach places and cooperative physical layer security with energy harvesting provides the secure signal transmission with improved coverage and unbreakable connection between the headquarters and the

soldiers. Further, the present era has completely shifted towards the wireless technologies, where there is a huge demand for low cost devices. With the help of RF EH, it is possible to obtain energy from the RF signals, resulting in efficient use of available energy. RF EH techniques are used in wireless body networks, wireless sensor networks, wireless-charging systems etc. Several commercial products are also available in the market based on the EH technique such as power caster and cota system. Due to various advantages of RF EH techniques, it can be collaborated with several types of networks such as cognitive radio networks (CRNs), heterogeneous networks and cellular networks for improving the energy efficiency of the system.

Physical Layer Security Model

The system model for the conventional system is shown in Figure 1. It consists of a source S, a destination D, a relay R, in presence of an eavesdropper E. The complex channel gains from S to R, from S to E, from R to D, and from R to E are denoted by h^*_{SR}, h^*_{SE}, h^*_{RD}, and h^*_{RE}, respectively. For the EH technique, the system model is shown in Figure 2. It is similar as that of the conventional system, except the beacon node, which is absent in the conventional system. Let h^*_{BR} and h^*_{BS} represents the complex channel gains from B to R and from B to S, respectively. It is assumed that the noise is additive white Gaussian noise (AWGN) with mean zero and variance σ^2 and self-interference is perfectly cancelled. Moreover, the relay switches between half and full duplex operations (Sinha, & Jindal, 2016). Full duplex relay (FDR) is able to receive and transmit signal at the same time.

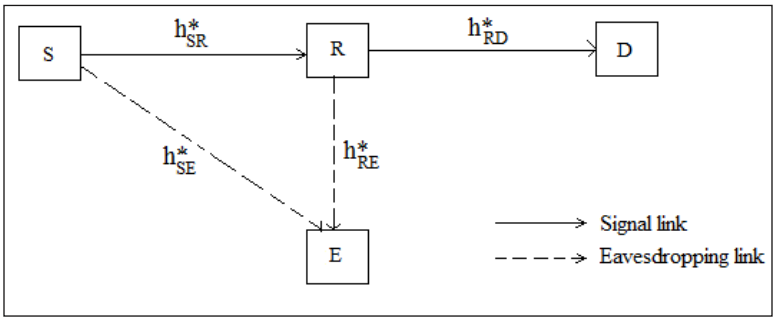


Figure 1. System model with single relay of conventional system

Energy Harvesting Protocol

In literature there are mainly three protocols used for EH system, TSR, PSR, IRR (ideal relay receiver).Among these TSR and PSR are most popularly used to replenish energy to deprived nodes during transmission.In two-phase communication the total time allotted to relay is divided between power transfer (PT) and information transmission (IT) time in TSR. On the other hand, total power is divided into two portions depending on power

splitting ration in PSR for PT and IT. Therefore, time allocation between PT and IT must be carefully designed to improve the throughput.In this paper, two energy harvesting protocols including time switching and power splitting have been analyzed.

A wireless communication system is developed, where legitimate nodes i.e. source S and destination D share confidential information with the aid of intermediate node. Beacon node B is used to allocate power to S and R for harvesting energy. The model has been implemented with the assumptions that there is no direct communication between S and D, relay node assist in between. Only one relay node is considered for simplicity. Relay and source nodes are energy constrained nodes and these nodes first harvest energy from beacon node and then use the harvested energy using TSR and PSR protocol. Amongst various relaying protocols, AF and DF are used and compared based on secrecy rate. Further, it is assumed that channel state information (CSI) is available.

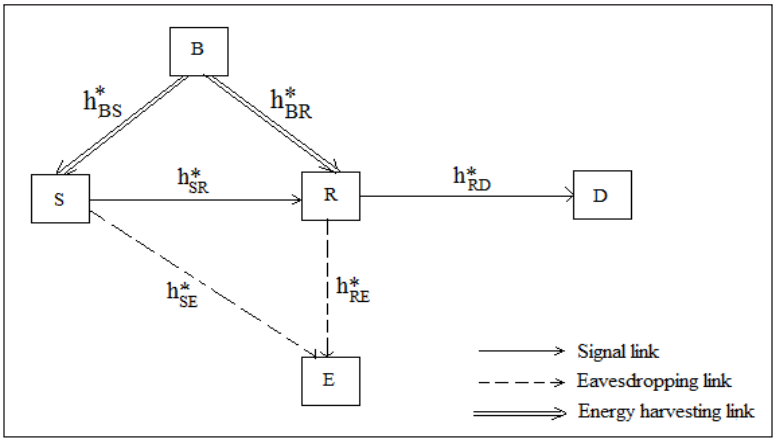


Figure 2. System model with single relay employing EH

Time Switching EH Protocol

Time switching protocol is shown in Figure 3, where T represents the total time required to transmit certain block of information from source to destination. As shown in Figure 3, the source and relay harvests energy from the beacon for duration of αT and $0 \leq \alpha \leq 1$. Data is transmitted from source to relay for half of the remaining time, and from relay to destination in remaining time duration (Sinha, & Jindal, 2017).

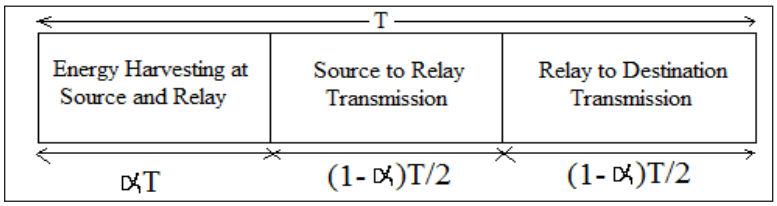


Figure 3. Time Switching EH Protocol

The harvested energy by S and R is given by (Nguyen, Duong, Ngo, Hadzi-Velkov, & Shu, 2016):

$$E_S = \eta P_B \alpha T \left| h_{BS}^* \right|^2 \quad [1]$$

$$E_R = \eta P_B \alpha T \left| h_{BR}^* \right|^2 \quad [2]$$

where, η represents efficiency of energy conversion technique, P_B denotes power of the beacon and α is the time fraction.

Therefore, the transmit power of S and R is given by (Nguyen, Duong, Ngo, Hadzi-Velkov, & Shu, 2016).

$$P_S = \frac{2\eta P_B \left| h_{BS}^* \right|^2 \alpha}{1-\alpha} \quad [3]$$

$$P_R = \frac{2\eta P_B \left| h_{BR}^* \right|^2 \alpha}{1-\alpha} \quad [4]$$

Power Splitting EH Protocol

Figure 4. showsPS EH protocol. As shown in the Figure 4, T represents time required to send a given block of information from sender to intended receiver. This duration is divided into two parts. During the first half, energy is harvested by source and relay, a portion of the signal is utilized for EH and rest of the signal is utilized for communication between source and relay. During the next half, relay sends the data to intended destination.

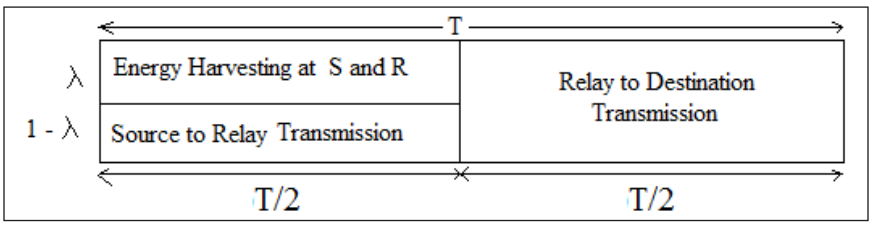


Figure 4. Power Splitting EH Protocol

The harvested energy by S and R is given by (Nasir, Zhou, Durrani, & Kennedy, 2013) as:

$$E_S = \eta \lambda P_B \left| h_{BS}^* \right|^2 (T/2) \quad [5]$$

$$E_R = \eta \lambda P_B \left| h_{BR}^* \right|^2 (T/2) \quad [6]$$

where, λ is the fraction of the signal used for EH.

Therefore, power transmitted by S and R is given by (Nasir, Zhou, Durrani, & Kennedy, 2013) as:

$$P_S = \eta\lambda P_B \left|h_{BS}^*\right|^2 \quad [7]$$

$$P_R = \eta\lambda P_B \left|h_{BR}^*\right|^2 \quad [8]$$

Cooperative Schemes

In this paper, two cooperation schemes amplify-and-forward and decode-and-forward have been analyzed with TS and PS EH protocols. The signal transmission with AF and DF schemes are detailed below:

Decode-and-Forward (DF) Scheme

It consists of two stages. The first stage involves transmission of information signal *x(n)* from source to relay and transmission of jamming signal *q(n)* by relay to the eavesdropper, simultaneously as shown in Figure 5.

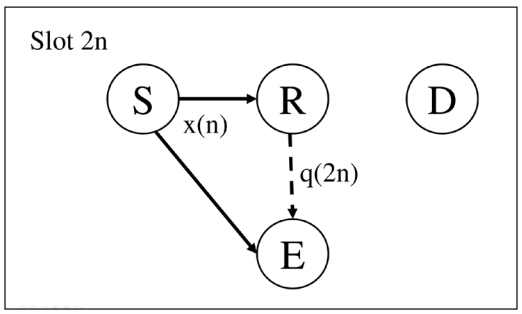


Figure 5. Signals transmitted in 2nth time slot

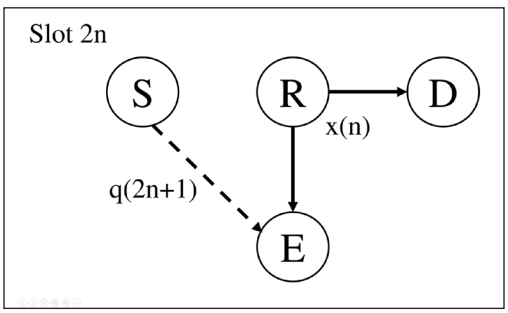


Figure 6. Signals transmitted in (2n+1)th time slot

Signals received by R and E during 2nth time slot are given by (Lee, 2015) as:

$$y_R(2n) = \sqrt{\rho P_S}\, h_{SR}^* x(n) + n_R(2n),$$

$$y_E(2n) \aleph \sqrt{\rho P_S}\, h_{SE}^* x(n) \quad \sqrt{\rho P_{RJ}}\, h_{RE}^* q(2n) \quad n_E(2n) \quad [9]$$

where, P_{RJ} denotes the jamming power of relay and $n_R(2n)$ and $n_E(2n)$ represents AWGN at relay and eavesdropper. Also, $\rho = 1$, for TS EH and $\rho = 1 - \lambda$, for PS EH.

In the second stage, relay first decode the received encoded signal, then re-encode it, and finally, transmits the re-encoded signal to destination. At the same time, source sends the jamming signal to the eavesdropper as shown in Figure 6.

Signals received by E and D in (2n+1)th time slot are given by (Lee, 2015):

$$y_E(2n+1) = \sqrt{P_R}\, h_{RE}^{*} x(n) + \sqrt{P_{SJ}}\, h_{SE}^{*} q(2n+1) + n_E(2n+1),$$
$$y_D(2n+1) = \sqrt{P_R}\, h_{RD}^{*} x(n) + n_D(2n+1) \quad [10]$$

where, P_{SJ} represents the jamming signal power of the source and n_D (2n + 1)denotes AWGN at D.

Amplify-and-Forward (AF) Scheme

This scheme also consists of two stages similar to that of DF scheme. Stage 1 is same as that of the DF scheme. Signals obtained at R and E during first stage is given by [9]. In second stage, relay sends amplified version of signal received to the destination. At the same instant, source sends jamming signal to destination. Signals received by D and E in time slot (2n+1) are given by:

$$y_D(2n+1) = G\sqrt{P_S}\, h_{RD}^{*} y_R(2n) + n_D(2n+1),$$
$$y_E(2n+1) = G\sqrt{P_S}\, h_{RE}^{*} y_R(2n) + \sqrt{P_{SJ}}\, h_{SE}^{*} q(2n+1) + n_E(2n+1) \quad [11]$$

Where, scaling factor (Kumar & Bhatia, 2015) is represented by $G = \frac{1}{\sqrt{P_S |h_{SR}|^2 + N_o}}$ and N_o represents variance of noise.

Achievable Secrecy Rate

Performance analysis has been done in terms of secrecy rate. Secrecy rate is defined as the amount of information that can be securely transmitted over the wireless medium in the presence of eavesdropper. The achievable secrecy rate with both AF and DF cooperation schemes is given follows:

DF Scheme. Using equations [9] and [10], the rates at *D* and *E* is given by

$$R_d = \frac{1}{2}\log_2(1 + P_R \alpha_{RD}) \quad [12]$$

$$R_e \aleph \frac{1}{\ddot{u}} \log_2\left(1 \quad \frac{\rho P_{\ddot{u}} \alpha}{+ \aleph P_{RJ} \quad _{RE}} \quad \frac{P_{\ddot{u}} \alpha}{+ P_{SJ} \quad _{SE}}\right) \quad [13]$$

where, $\alpha_{RD} = \frac{|h_{RD}|^2}{\sigma^2}$, $\alpha_{SE} = \frac{|h_{SE}|^2}{\sigma^2}$ and $\alpha_{RE} = \frac{|h_{RE}|^2}{\sigma^2}$. Using equations [12] and [13], the secrecy rate is given by $R_S = \max\{R_d - R_e, 0\}$, where

$$R_d - R_e = \frac{1}{2}\log_2\left(\frac{1+P_R\alpha_{RD}}{1+\frac{\rho P_S\alpha_{SE}}{1+\rho P_{RJ}\alpha_{RE}}+\frac{P_R\alpha_{RE}}{1+P_{SJ}\alpha_{SE}}}\right) \quad [14]$$

AF Scheme. Using equations [9] and [11], the rates at D and E are represented as in (Lee, 2015):

$$R_d = \frac{1}{2}\log_2(1+G^2P_S\alpha_{RD}) \quad [15]$$

$$R_e = \frac{1}{2}\log_2\left(1+\frac{\rho P_S\alpha_{SE}}{1+\rho P_{RJ}\alpha_{RE}}+\frac{G^2P_S\alpha_{RE}}{1+P_{SJ}\alpha_{SE}}\right) \quad [16]$$

The secrecy rate is given by $R_S = \max\{R_d - R_e, 0\}$, where

$$R_d - R_e = \frac{1}{2}\log_2\left(\frac{1+G^2P_S\alpha_{RD}}{1+\frac{\rho P_S\alpha_{SE}}{1+\rho P_{RJ}\alpha_{RE}}+\frac{G^2P_S\alpha_{RE}}{1+P_{SJ}\alpha_{SE}}}\right) \quad [17]$$

RESULTS AND DISCUSSION

Numerical results are presented for investigating secrecy rate of system utilizing both TS and PS EH using AF and DF relaying techniques. The performance of the system employing EH is compared with the conventional system in terms of secrecy rate. It is assumed that the source *S*, destination *D* and relay *R* and are located in a line (Sinha, & Jindal, 2016) as shown in Figure 7, where, d_{BS}, d_{BR}, d_{SR}, d_{RE} and d_{RD} show *B*-*S*, *B*-*R*, *S*-*R*, *R*-*E* and *R*-*D* distance. The distance between *E* and *S* can be represented as $d_{SE} = \sqrt{d_{SR}^2 + d_{RE}^2}$, respectively. The channel used between the nodes is the line-of-sight (LOS) channel $d^{-\frac{c}{2}}e^{j\theta}$, where *d* represents node to node distance, θ is the phase that is having uniform distribution within [0, 2π, and the path loss exponent is given by c=3.5 (Dong, Han, Petropulu, & Poor, 2010).

The performance of EH system is affected by many factors like, distance between relay and eavesdropper node, distance between relay and destination node, beacon power etc. Keeping these factors into consideration, the simulation is conducted to analyze the performance of communication system. The figure of merit of this work is secrecy rate comparison of conventional and EH-TSR, PSR system. All the nodes in experimental setup are deployed to investigate the system performance as shown in Table 1.

Table 1
Distance between communication nodes in all cases of experimental setup

Cases	Conventional and EH system					
	$\mathbf{d_{SR}}$	$\mathbf{d_{RE}}$	$\mathbf{d_{RD}}$	$\mathbf{d_{SE}}$	$\mathbf{d_{BS}}$	$\mathbf{d_{BR}}$
I	10m	15-40m	15m	18.02m-41.23m	7m	7m
II	10m	15m	5-30m	18.02m	7m	7m
III	10m	15m	15m	18.02m	7m	7m
IV	10m	15m	15m	18.02m	6-15m	6-15m

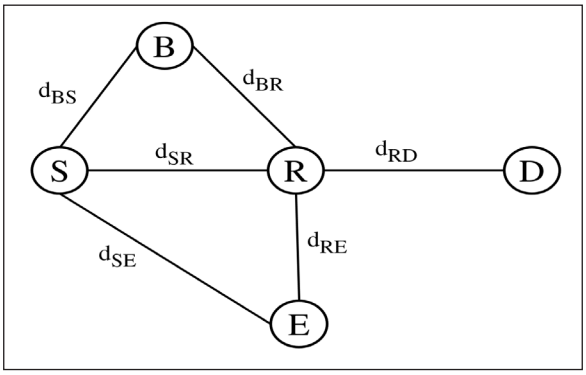


Figure 7. Simulation Model

For the conventional system, it is assumed that the system functions under the constraint of total transmit power of 30 dBm and noise power is -70 dBm. For the system with EH, it is assumed that P_B = 30 dBm, the noise power = -70 dBm and d_{BS}= d_{BR}= 7 m. Further, η =1 for both EH schemes. For TS EH scheme, it is assumed that α=0.999 and for PS EH, it is assumed that λ =0.99. Moreover, the nodes utilize equal power allocation scheme.

The conventional system and EH are investigated for both TSR and PSR protocol using AF and DF methods. To evaluate the performance of wireless system there are many parameters such as secrecy rate, throughput, secrecy outage probability (SOP), ergodic secrecy capacity (ESC) and secrecy rate used in the literature. Here, secrecy rate is considered as parameter metric as it is a measure to ensure the authentication, confidentiality and integrity during data transmission without increased complexity.

The following cases depict various simulation results:

CASE 1: Secrecy Rate versus Relay-Eavesdropper Distance

(a) For Conventional System: Figure 8.shows the plot of secrecy rate versus relay-eavesdropper distance for the conventional system using AF and DF schemes. Secrecy rate increases with increase in distance between relay and eavesdropper.
(b) For TS EH System: Figure 9.shows the plot of secrecy rate versus relay eavesdropper

distance for the system employing TS EH for AF and DF techniques. The plot indicates that with increase in distance between relay and eavesdropper, secrecy rate also increases.

(c) For PS EH System: Figure 10 shows the plot of secrecy rate versus relay-eavesdropper distance for system employing PS EH for AF and DF cooperative schemes. This plot also shows that as the eavesdropper shifts away from the relay, the transmission becomes more secure.

The numerical values for the variation in secrecy rate when the distance between the relay and eavesdropper is 40m have been observed. It is observed that secrecy rate for TS EH is 11% more but for PS EH is 44% less as compared to conventional system (without applying EH technique) with AF cooperative scheme. Further, secrecy rate for TS EH is 11% more but for PS EH is 52% less as compared to conventional system for DF cooperative scheme. The secrecy rate obtained with TS EH is 44% and 57% more as compared to PS EH in AF and DF respectively. Further, the secrecy rate obtained for the conventional physical layer security system with AF is 8.16% more as compared to system with DF.

CASE 2: Secrecy Rate versus Relay-Destination Distance

(a) For Conventional System: Figure 11 shows the plot of secrecy rate versus relay destination distance for the conventional system using AF and DF cooperative schemes. The plot shows that, with increase in the distance between relay and destination, the secrecy rate decreases.

(b) For TS EH System: Figure 12 shows the plot of secrecy rates versus relay-destination distance for the system employing TS EH. The plot indicates that secrecy rate decreases as the distance between the relay and destination increases.

(c) For PS EH System: The plot of secrecy rates versus relay-destination distance for the system employing PS EH is shown in Figure 13. The plot shows that, secrecy rate decreases with increase in distance between relay and destination.

CASE 3: Secrecy Rate versus Path Loss Exponent

(a) For Conventional System: Figure 14 represents the plot of secrecy rate versus path loss exponent for conventional system. It is indicated by the plot that with increase in the path loss exponent, channel become worse and hence the transmission becomes less secure.

(b) For TS EH System: Figure 15 represents the plot of secrecy rate with respect to path loss exponent for system employing TS EH for DF and AF cooperative schemes. The plot shows that secrecy rate decreases with increase in the path loss exponent.

(c) For PS EH System: Figure 16 shows the plot of secrecy rate with respect to path loss exponent for PS EH system. The communication becomes less secure with increase in the pass loss exponent.

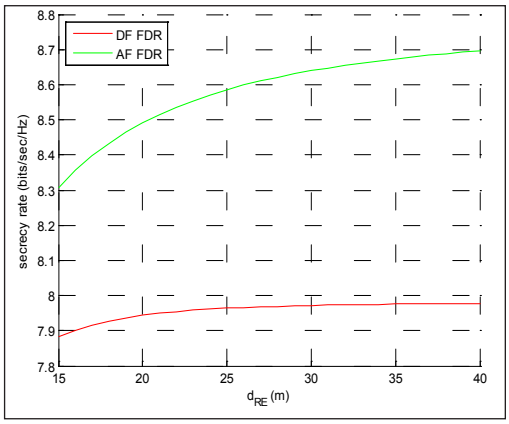


Figure 8. Secrecy rate with respect to d_{RE} for conventional system without EH, where d_{SR}= 10 m, d_{RD}= 15 m

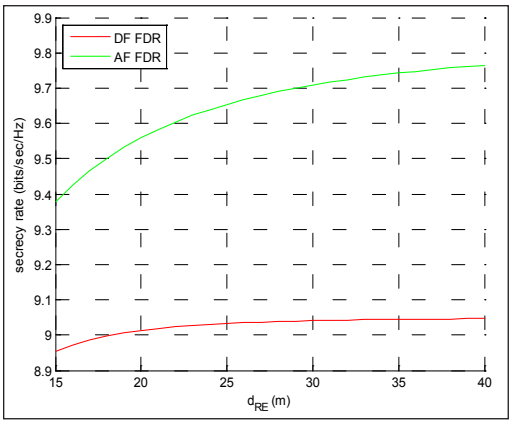


Figure 9. Secrecy rate with respect to d_{RE} for system with TS EH, where d_{SR}= 10 m, d_{RD}= 15 m, d_{BS}= 7 m and d_{BR} = 7 m

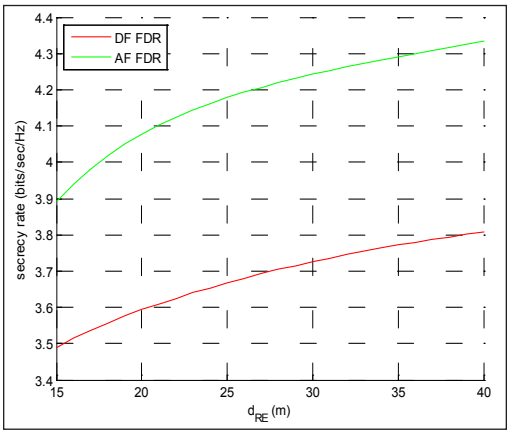


Figure 10. Secrecy rate with respect to d_{RE} for system with PS EH, where d_{SR}= 10 m, d_{RD}= 15 m, d_{BS}= 7 m and d_{BR} = 7 m

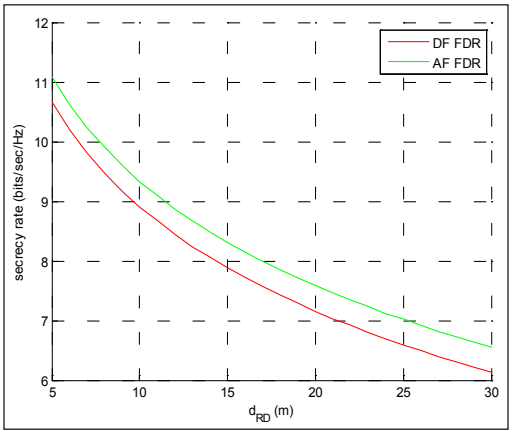


Figure 11. Secrecy rate with respect to d_{RD} for conventional system without EH, where d_{SR}= 10 m, d_{RE}= 15 m

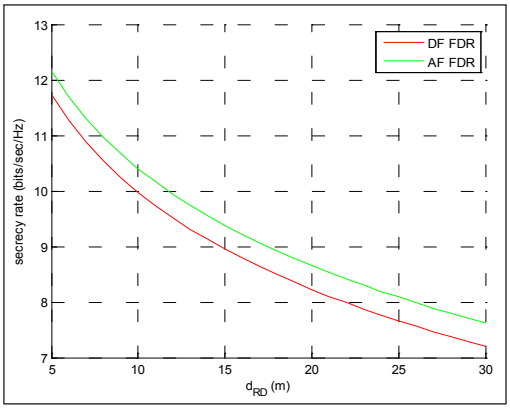


Figure 12. Secrecy rate with respect to d_{RD} for system with TS EH, where d_{SR}= 10 m, d_{RE}= 15 m d_{BS}= 7 m and d_{BR} = 7 m

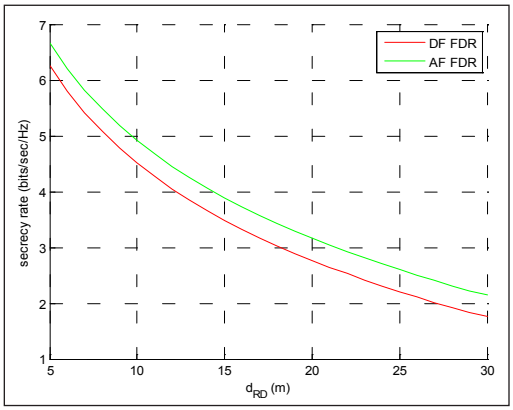


Figure 13. Secrecy rate with respect to d_{RD} for system with PS EH, where d_{SR}= 10 m, d_{RE}= 15 m d_{BS}= 7 m and d_{BR} = 7 m

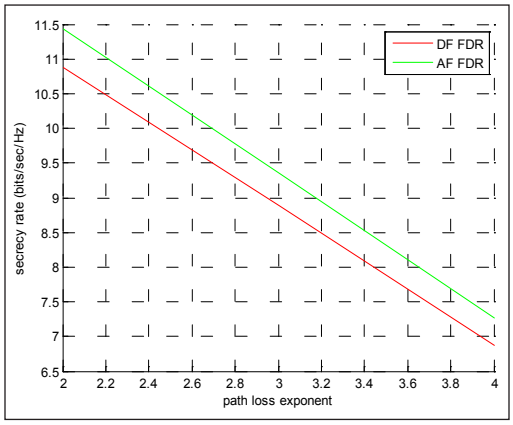


Figure 14. Secrecy rate with respect to path loss component for conventional system, where d_{SR}= 10 m, d_{RE}= 15 m, d_{RD}= 15 m

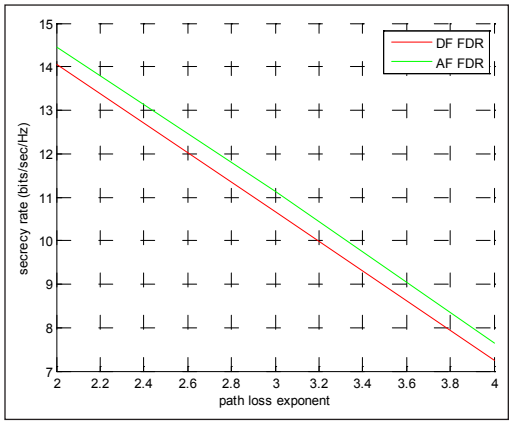


Figure 15. Secrecy rate with respect to path loss exponent for system with TS EH, where d_{SR}= 10 m, d_{RE}= 15 m, d_{RD}= 15 m, d_{BS}= 7 m and d_{BR} = 7 m

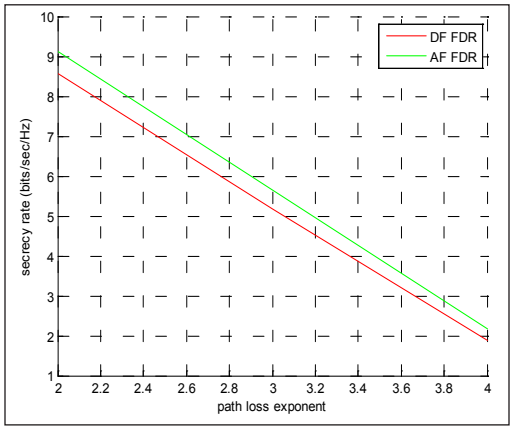


Figure 16. Secrecy rate with respect to path loss exponent for system with PS EH, where d_{SR}= 10 m, d_{RE}= 15 m, d_{RD}= 15 m, d_{BS}= 7 m and d_{BR} = 7 m

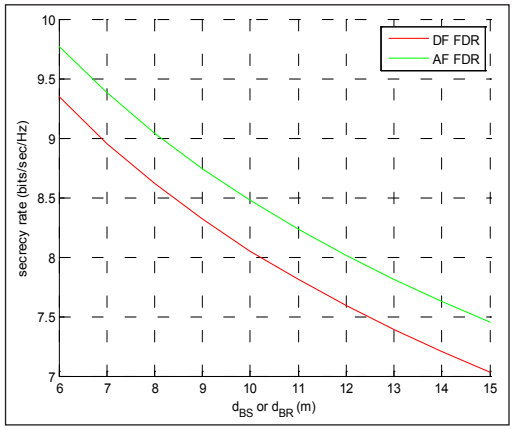


Figure 17. Secrecy rate versus d_{BS}or d_{BR} when d_{SR}= 10 m, d_{RE}= 15 m, d_{RD}= 15 m in system with TS EH

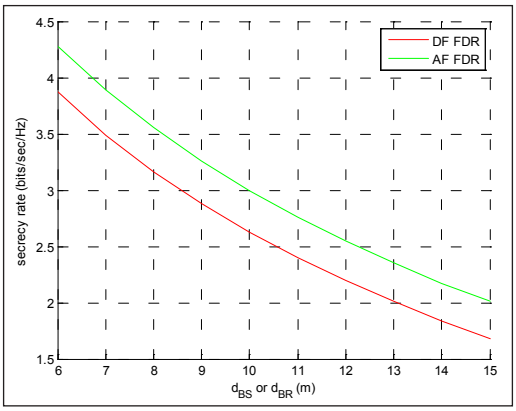


Figure 18. Secrecy rate with respect to d_{BS}or d_{BR} for system with PS EH, where d_{SR}= 10 m, d_{RE}= 15 m, d_{RD}= 15 m

CASE 4: Secrecy Rate versus Beacon Distance

(a) For TS EH System: Figure 17 shows the plot of secrecy rate versus beacon-source or beacon-relay distance in the system with TS EH. The plot shows that secrecy rate decreases as the beacon moves away from the source and relay.

(b) For PS EH System: The plot of secrecy rate versus beacon-source or beacon-relay distance in the system with PS EH is shown in Figure 18. It is represented in the plot that when the beacon moves away from source and relay, secrecy rate decreases.

All the plots clearly show that the performance of the system with TS EH is better than that of the conventional system and the performance with conventional system is better than that of the system with PS EH in the given condition. The reason is that, in case of TS EH system the whole signal strength is used by the nodes for harvesting energy and transmission of information, whereas in case of PS EH, the signal strength is divided into two parts, one for energy harvesting and the other for data transmission. Moreover, it has been observed that, in all network scenarios, the AF cooperative scheme outperforms the DF cooperative scheme.

CONCLUSION

In this paper, we have studied and presented a deep insight of the combination of three technologies; energy harvesting, physical layer security and cooperative communication. The work presented in this paper demonstrates that how energy harvesting along with providing secure communication prolonged the lifetime of energy constrained devices. We have observed the numerical values for the variation in secrecy rate when the distance between the relay and eavesdropper is 40m. It is observed that secrecy rate for TS EH is 11% more but for PS EH is 44% less as compared to conventional system (without applying EH technique) with AF cooperative scheme. Further, secrecy rate for TS EH is 11% more but for PS EH is 52% less as compared to conventional system for DF cooperative scheme. The secrecy rate obtained with TS EH is 44% and 57% more as compared to PS EH in AF and DF respectively. Further, the secrecy rate obtained for the conventional physical layer security system with AF is 8.16% more as compared to system with DF. The performance analysis presented herein may be used as reference for selecting the particular energy harvesting protocol with cooperative scheme for given applications as required. An in-depth comparative analysis of different energy harvesting protocols with physical layer security in different cooperative scenarios makes this paper different from the existing work as the existing literature does not present such an extensive comparative analysis of the problem under study.

Further, implementation of RF-EH in multi-user environment, implementations of multi relay nodes to select an efficient relay and jammer algorithms for improved system performance, implementation of multiple antennas at relay node for improving transmission efficiency, redesigning of conventional schemes and protocols to optimise power consumption in existing protocols while maintaining secrecy constraints of the cooperative network are some open research problems in the area under study.

REFERENCES

Chen, G., Gong, Y., Xiao, P., & Chambers, J. A. (2015). Physical layer network security in the full-duplex relay system. *IEEE Transactions on Information Forensics and Security*, *10*(3), 574-583.

Dong, L., Han, Z., Petropulu, A. P., & Poor, H. V. (2010). Improving wireless physical layer security via cooperating relays. *IEEE Transactions on Signal Processing*, *58*(3), 1875-1888.

Hoang, T. M., Duong, T. Q., VO, N. S., & Kundu, C. (2017). Physical layer security in cooperative energy harvesting networks with a friendly jammer. *IEEE Wireless Communications Letters*, *6*(2), 174-177.

Jindal, P., & Sinha, R. (2017, March). Physical layer security with energy harvesting in single hop wireless relaying system. In *International Conference on Information Science and Applications* (pp. 249-256). Singapore: Springer.

Kumar, N., & Bhatia, V. (2015). Performance analysis of amplify-and-forward cooperative networks with best-relay selection over Weibull fading channels. *Wireless Personal Communications*, *85*(3), 641-653.

Lee, J. H. (2015). Full-duplex relay for enhancing physical layer security in multi-hop relaying systems. *IEEE Communications Letters*, *19*(4), 525-528.

Liu, W., Zhou, X., Durrani, S., & Popovski, P. (2016). Secure communication with a wireless-powered friendly jammer. *IEEE Transactions on Wireless Communications*, *15*(1), 401-415.

Nasir, A. A., Zhou, X., Durrani, S., & Kennedy, R. A. (2013). Relaying protocols for wireless energy harvesting and information processing. *IEEE Transactions on Wireless Communications*, *12*(7), 3622-3636.

Nguyen, N. P., Duong, T. Q., Ngo, H. Q., Hadzi-Velkov, Z., & Shu, L. (2016). Secure 5G wireless communications: A joint relay selection and wireless power transfer approach. *IEEE Access*, *4*, 3349-3359.

Salem, A., Hamdi, K. A., & Rabie, K. M. (2016). Physical layer security with RF energy harvesting in AF multi-antenna relaying networks. *IEEE Transactions on Communications*, *64*(7), 3025-3038.

Sinha, R., & Jindal, P. (2016, November). Performance analysis of cooperative schemes under total transmit power constraint in single hop wireless relaying system. *2nd International Conference* in *Communication Control and Intelligent Systems (CCIS), 2016* (pp. 28-31). Mathura, India.

Wyner, A. D. (1975). The wire-tap channel. *Bell Labs Technical Journal*, *54*(8), 1355-1387.

Yuen, C., Elkashlan, M., Qian, Y., Duong, T. Q., Shu, L., & Schmidt, F. (2015). Energy harvesting communications: Part 2 [Guest Editorial]. *IEEE Communications Magazine*, *53*(6), 54-55.

Yuen, C., Elkashlan, M., Qian, Y., Duong, T. Q., Shu, L., & Schmidt, F. (2015). Energy harvesting communications: Part 1 [guest editorial]. *IEEE Communications Magazine*, *53*(4), 68-69.

Yuen, C., Elkashlan, M., Qian, Y., Duong, T. Q., Shu, L., & Schmidt, F. (2015). Energy harvesting communications: Part III [Guest Editorial]. *IEEE Communications Magazine*, *53*(8), 90-91.

Zhang, C., & Chen, Y. (2017). Wireless Power Transfer Strategies for Cooperative Relay System to Maximize Information Throughput. *IEEE Access*, *5*, 2573-2582.

Zhang, R., & Ho, C. K. (2013). MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Transactions on Wireless Communications*, *12*(5), 1989-2001.

Son, P. N., & Kong, H. Y. (2015). Cooperative communication with energy-harvesting relays under physical layer security. *IET Communications*, *9*(17), 2131-2139.

Wu, W., Wang, B., Deng, Z., & Zhang, H. (2017). Secure beamforming for full-duplex wireless powered communication systems with self-energy recycling. *IEEE Wireless Communication Letters, 6*(2), 146-149.

Pertanika J. Sci. & Technol. 26 (4): 1693 - 1712 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Feasibility Study Using Electrolysis Treatment (ET) As the Pre-treatment Method to Extract Lipid from *Chlorella* sp. for Biodiesel Production

Rachel Fran Mansa[1]*, Coswald Stephen Sipaut[1], Suhaimi Md.Yasir[2], Jedol Dayou[3] and Costantine Joannes[1]

[1]Energy and Materials Research Group (EMRG), Faculty of Engineering, Universiti Malaysia Sabah (UMS), Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia
[2]Seaweed Research Unit (UPRL), Faculty of Science and Natural Resources, Universiti Malaysia Sabah (UMS), Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia
[3]Energy, Vibration and Sound Research Group (e-VIBS), Faculty of Science and Natural Resources, Universiti Malaysia Sabah (UMS), Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia

ABSTRACT

The feasibility study on *Chlorella* sp. lipid extraction using an electrolysis treatment (ET) as pre-treatment was investigated. Stainless steel was used as the anode and cathode material. The ET method was conducted in a batch or continuous system with or without air aeration and recycling flow. The total lipid in *Chlorella* sp. A_{WET} and A_{WET} were not analysed due to small sample volume. Approximately same amount of lipids were attained from *Chlorella* sp. B_{WOET} ($7.97 \pm 0.43\%$ $g_{lipid}/g_{dry\ wt}$) and B_{WET} ($7.95 \pm 0.37\%$ $g_{lipid}/g_{dry\ wt}$) if treated at 5 V/cm and aerated at 16.7 $\mu m^3/s$ for 1800s. Whereas, if *Chlorella* sp. was treated at 13 V/cm and aerated at 16.7 $\mu m^3/s$ for 1800 s, the total lipid obtained in *Chlorella* sp. C_{WOET} ($8.18 \pm 0.49\%$ $g_{lipid}/g_{dry\ wt}$) was 1.13-fold higher than C_{WET} ($7.22 \pm 0.47\%$ $g_{lipid}/g_{dry\ wt}$). Meanwhile under semi-continuous system, similar pattern of result was achieved in *Chlorella* sp. D_{WOET} ($8.58 \pm 0.49\%$ $g_{lipid}/g_{dry\ wt}$) with 1.11-fold higher than D_{WET} ($7.72 \pm 0.54\%$ $g_{lipid}/g_{dry\ wt}$), if treated at 14 V/cm and recycled at 2.3 $\mu m^3/s$ for 3000s. This corresponded to lipid oxidation that might have occurred during the ET method. The fatty acid methyl ester (FAME) composition of *Chlorella* sp. D_{WOET} and D_{WET} contained predominantly methyl linolenate (C18:3) and methyl palmitate

ARTICLE INFO

Article history:
Received: 27 August 2017
Accepted: 21 June 2018
Published: 24 October 2018

E-mail addresses:
rfmansa@ums.edu.my (Rachel Fran Mansa)
css@ums.edu.my (Coswald Stephen Sipaut)
ysuhaimi@ums.edu.my (Suhaimi Md. Yasir)
jed@ums.edu.my (Jedol Dayou)
costantinejoannes@gmail.com (Costantine Joannes)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

(C16:0). The concentrations of methyl palmitate attained in *Chlorella* sp. D_{WOET} and D_{WET} were 0.049 ± 0.005 g/m^3 and 0.045 ± 0.005 g/m^3, respectively.

Keywords: *Chlorella* sp., electrolysis treatment, lipid extraction, methyl palmitate, total lipid

INTRODUCTION

This depletion of fossil fuel, in addition to fluctuations in oil prices and global warming issues, makes the development of alternative and renewable resources of energy crucial – a need reflected in a 6.5% increase in biodiesel production worldwide in 2016 (BP Statistical Review of World Energy, 2017). Biodiesel from microalga has received considerable attention over the last decade. Microalgae have many desirable features: a flexibility to grow in various conditions, a high growth rate, lipid-rich content, non-seasonal production type, non-arable land requirement, and non-food crop status. These attractive properties have driven interest in the use of microalgae as a substitute for conventional biodiesel feedstock (Japar et al., 2017; Kumar & Sharma, 2016; Patel et al., 2017; Suali & Sarbatly, 2012).

The process of producing biodiesel from microalgae involves microalgae cultivation, microalgae harvesting, microalgae lipid extraction and transesterification of lipid to fatty acid methyl esters (FAME) (Japar et al., 2017; Mansa et al., 2012; Velasquez-Orta et al., 2013). Of all the processes, lipid extraction is one of the most complex and challenging. Conventional extraction technologies such as high pressure homogenizing, supercritical carbon dioxide extraction, sonication, microwave extraction, osmotic shock, solvent extraction and enzymatic degradation are commonly used in microalgae lipid extraction (Chen et al., 2016; Jeevan-Kumar et al., 2017). Some researchers have also conducted a pre-treatment based on mechanical or physical extraction method as a purpose to disrupt the microalgae rigid cell wall and followed by the solvent extraction method for easy access of the microalga lipid (Joannes et al., 2015a). These methods are found to be efficient but may have several drawbacks such as high energy and time consumption, high toxicity or flammability, expensive raw materials, and difficulty in scaling up (Günerken et al., 2015).

A recent development in extraction technology using electric fields to disrupt the cell wall of microalgae has drawn interest for the last decade such as pulsed electric field (PEF) (Grimi et al., 2014; Lam et al., 2017; Parniakov et al., 2015; Postma et al., 2016). The microalga lipid extraction can be extracted via PEF alone and improved using PEF prior to solvent extraction method (Joannes et al., 2015a). Although, the PEF extraction process is conducted in a very short time period it involves a very high voltage (0.075 - 45 kV) (Joannes et al., 2015b). Therefore, extra precautions are mandatory if using PEF. Aside from PEF, another new electrical based approach to microalgae lipid extraction is the use of a direct current (DC) at a lower voltage such as electrolysis treatment (ET) method (Daghrir et al., 2014; Guldhe et al., 2016; Hua et al., 2016; Misra et al., 2015, 2014). The method is conducted under a batch system in microalgae harvesting stage (Guldhe et al., 2016;

Misra et al., 2015, 2014; Zhou et al., 2016). Moreover, electrolytes such as sodium chloride and sodium sulphate are usually being added into the sample to enhance its conductivity.

As described by Uduman et al. (2010) the electrodes with the same or different materials are submerged into a container which consist of microalga and then the electric is applied. This creates a constant electrical potential difference between the positive (anode) and negative terminal (cathode). The microalga cells (negatively charged) will tend to coagulate at anode (Uduman et al., 2010). This phenomenon also called as electrophoresis or the movement of small, suspended particles in a liquid via potential differences between the electrodes (Barros et al., 2015). In other theory, similar to PEF the electric field induction will causes a phenomenon called reversible or an irreversible electroporation of the cell membrane (Dimitrov, 1995; Kotnik et al., 2012). It is believed that, every cell membrane has its own critical strength. When the electric fields strength is increased above the critical strength of the cell membrane, the forces of noncovalent interaction that binds the lipid bilayer together will be overcome by the electric field forces (Kotnik et al., 2012). This will then lead the formation of small pores or holes on the surface of the cell membrane. Exposing the cell to very high field strength will cause the enlargement of pores and enhance the pores number. Thus, resulting in the released of intracellular components such as lipid, proteins, carbohydrates.

Hence**,** the aim of this work is to investigate the feasibility of using the electrolysis treatment (ET) method and without the addition of any electrolyte to extract microalgae lipid for biodiesel production. This work is used a direct electrolysis (without any additional of electrolytes) specifically for microalgae lipid extraction which has not yet been reported so far. Instead of using non-sacrificial carbon electrodes (Guldhe et al., 2016; Misra et al., 2014, 2015) the present study used stainless steel material for both cathode and anode. Herein, *Chlorella* sp. was pre-treated using ET method prior to solvent extraction. It was expected that, by using ET, the lipid extraction can be improved or enhanced.

MATERIALS AND METHODS

Microalga Strain, Media and Culture Condition

Local freshwater microalga *Chlorella* sp. was obtained from the Borneo Marine Research Institute (BMRI), Universiti Malaysia Sabah (Joannes et al., 2016). The cell sizes were in the range 3 – 8 µm. The microalga was cultivated in a 0.002 m^3 Erlenmeyer flask with five batches for 30 days of cultivation under photoautotrophic mode using a Bold's Basal Medium (BBM). The cultivation process was conducted under a florescent white cool lamp with light intensity of 24 $\mu mol/m^2 \cdot s$ and aerated through air bubble. The culture temperature and pH were maintained at 24.4 ± 0.1^oC and 8.89 ± 0.05^oC, respectively. Cultivation was conducted under a cycle of 12 h light and 12 h dark.

Microalga Cell Concentration

The microalga cell concentration was manually determined by dividing the microalga dried weight with its volume and can be referred in Eq. [1]. Approximately 20 μm^3 of microalga culture was taken using a pipette and was placed in a 25 μm^3 container. The initial mass of the container was weighed using an analytical balancer with an accuracy of ± 0.0001 g. Next, the sample was placed in an oven at 70°C and left overnight to dry. Before recording its final mass, the sample was placed in a desiccator and cooled at a room temperature for 2 to 3 mins. Then, the final mass of the container was recorded. The sample was placed into the oven again for an additional 30 mins and reweighed until a constant reading was obtained. The above steps were repeated twice. This was to ensure the final weight of the dried microalga was achieved.

$$\text{Microalga concentration} = \frac{\text{Dried weight (kg)}}{\text{Volume } (m^3)} \times 100\% \quad [1]$$

Experimental Parameters and Chamber Design Setup

The schematic diagram of the electrolysis treatment (ET) chamber was shown in Figure 1. The ET chamber was made of 0.0025 m thick glass (soda-lime-silica glass). Its dimensions were 0.174 m (length) × 0.12 m (height) × 0.051 m (width). Stainless steel with dimensions of 0.154 m (length) × 0.109 m (width) and 0.0004 m thickness was used as the electrode material in anode and cathode. The gap between the electrodes was fixed at 0.022 m. The length and inner diameter of the tube used in this study were 1.5 m and 0.004 m, respectively. The electrodes were connected to a DC power supply (MP 303-3, Meguro, Malaysia). A peristaltic pump was used to create the recycling flow at 2.3 $\mu m^3/s$ and was measured using a digital flowrate meter. The aeration flow was formed using a bubble stone at 16.7 $\mu m^3/s$. Volume of sample used were 60 μm^3 for A sample and 0.5 μm^3 for B, C, D, H_1, H_2 and H_3 samples. The microalga cell concentrations used were in the range 0.18 – 0.46 kg/m^3. Voltages from 11 to 31 V were applied to form an electric field between the

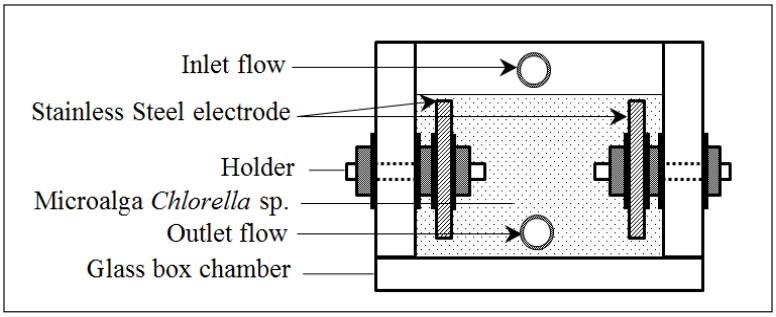


Figure 1. Schematic diagram of the ET chamber (front view)

electrodes of 5 – 14 V/cm. The experiment (A, B and C condition) was conducted under a batch system except for D condition which was conducted under a semi-continuous system. The experiment was conducted in duplicate and all values presented were in average ± standard errors unless stated.

Direct Heat Treatment Setup

A 0.0005 m^3 of *Chlorella* sp. labelled as H_1, H_2 and H_3 were placed in a beaker, stirred and heated at 30°C, 45°C and 60°C using a magnetic stirrer (HS0707V2, Favorit, Malaysia) for 1800s, respectively. A digital thermometer was also used to double check the temperature of the sample.

Lipid Extraction and Total Lipid Determination

After the ET method, the sample was left overnight to allow the formation of three layers: microalgae biomass at the bottom layer, water in the intermediate layer, and unknown components in the upper layer. *Chlorella* sp. without electrolysis treatment (WOET) and with electrolysis treatment (WET) samples were centrifuged at 8000 rpm for 15 mins, frozen at -78°C overnight, and then freeze-dried at -41 ˚C under vacuum conditions for 24 h. The total lipid was extracted from the lyophilized biomass via the Folch method (using a solvent mixture of 13.3 μm^3 of chloroform and 6.7 μm^3 of methanol, $CHCl_3$-MeOH; 2:1; v/v) (Folch et al., 1957) and left for 12 h without stirring. The biomass was filtered using filter paper and the filtrate was washed using 4 μm^3 of water to separate the $CHCl_3$ (bottom layer) and MeOH-water (upper layer). The bottom layer was extracted using a Pasteur pipette, evaporated at 62°C using an evaporator and the residue was dried at 70°C for 2 h. The sample was left to cool at room temperature and the final weight of the extract was recorded. The total lipid content was calculated using Eq. [2] (Abdullah et al., 2017).

$$\text{Total lipid (\%)} = \frac{\text{Weight of lipid (g)}}{\text{Weight of dried microalga (g)}} \times 100\% \qquad [2]$$

Scanning Electron Microscopy

Approximately 10 μm^3 of *Chlorella* sp. from D_{WOET} and D_{WET} samples were harvested and centrifuged at 8000 rpm for 5 mins. Then, the samples were frozen at -78°C overnight. The concentrated cells were lyophilized using a freeze dryer at -4°C under vacuum conditions for 24 h. Next, the surface morphologies of the samples were examined by a scanning electron microscopy (SEM) (EVO MA 10, Carl Zeiss, Germany) (Joannes et al., 2016). The sample was viewed under 1000× and 3000 × magnification powers.

Microalga Lipid Functional Group Determination

Approximately, 500 μm^3 of $CHCl_3$ (containing lipids) from bottom phase was collected using a 1 μm^3 micropipette and placed into a glass vial with cap. A Fourier transform infrared spectroscopy (FTIR) (2000 Series, Perkin-Elmer, UK) was used to determine the functional groups of lipid from D_{WOET} and D_{WET} samples. The wave number from 800 to 2600 cm^{-1} was set to scan the lipid functional groups.

FAME Conversion Procedures

The extracted lipid was converted to fatty acid methyl esters (FAME) using the method and procedures described by Zbinden et al. (2013). First, 0.5 μm^3 of a mixture of MeOH and toluene (1:1; v/v) was added to dissolve the dried lipid sample. Toluene was used to solubilize the non-polar lipids (e.g. cholesterol esters or triacylglycerol). Second, 0.5 μm^3 of 0.2 kg/m^3 methanol KOH solution was added. The sample was sealed in a 28 μm^3 universal bottle and vortexed. Next, it was heated at 37°C for 15 min in a water bath to allow the transesterification reaction to occur and then was cooled at room temperature for a few minutes. 0.5 μm^3 of acetic acid solution 0.2 kg/m^3 was added into the sample, followed by 2.0 μm^3 of $CHCl_3$ and 2.0 μm^3 of deionized water (DI). Mixing by swirling was performed for each addition of the reagents. The sample was then centrifuged at 1500 rpm for 5 mins. After the two layers were formed, the bottom layer ($CHCl_3$ phase) was extracted using a Pasteur pipette and was placed into a new and clean universal bottle. Finally, 1.0 μm^3 of $CHCl_3$ was added into the sample again for further purification and the sample was centrifuged once more, extracted, and placed into a new universal bottle. These steps were repeated twice. The extracted sample was to be further tested in a GCMS for FAME content.

GCMS Analysis Procedures

The analysis of FAME was performed using Gas chromatography mass spectrophotometer (GCMS) (Model 6890N, Agilent Technologies, USA). First, 1 μm^3 of the sample was injected into the GC capillary of 30.0 m × 0.25 mm × 0.25 µm ID nominal. The GCMS setting was adapted from Zbinden et al. (2013) but used zero holding time at maximum temperature. The flow rate of Helium gas (the carrier gas) was set at 1.4 μm^3/min with an initial temperature of 120°C held for 2 mins, then increased 6°C/min to 180°C, 1.5°C/min to 198°C, 5°C/min to 240°C. The split ratio was set at 120:1. The methyl palmitate (n-hexadecanoic acid methyl ester) standard was used as an external standard for FAME and was purchased from Sigma-Aldrich (≥ 99% purities). The methyl palmitate standard was prepared with five different concentrations ranging from 0.36 ± 0.04 $\mu g/m^3$ to 3.63 ± 0.06 $\mu g/m^3$. The concentrations of methyl palmitate in *Chlorella* sp. D_{WOET} and D_{WET}

samples were only determined and quantified using the standard curve of y = 18406792x ($R^2 = 0.995$); where y represents the peak area in cm^2 and x represents the methyl palmitate concentration in g/m^3.

RESULTS AND DISCUSSION

Total Lipid Content Extracted Based On Test Conditions

The total lipid content in each samples was quantified except from A_{WOET} and A_{WET} samples. This was due to sample volume used was too small (60 μm^3), hence not enough for further solvent extraction. From Table 1, it can be seen that the total lipid of 0.39 ± 0.02 kg/m^3 *Chlorella* sp. obtained from the B_{WOET} and B_{WET} were 7.97 ± 0.43% $g_{lipid}/g_{dry\ wt}$ and 7.95 ± 0.37% $g_{lipid}/g_{dry\ wt}$, respectively. These values were approximately the same. Meaning to say, using the ET method at 5 V/cm for 1800s and with an air aerated at 16.7 $\mu m^3/s$ was not able to improve the lipid extraction. The field strength applied might not be enough to weaken the *Chlorella* sp. cells wall. By maintaining the treatment time and aeration flow but then increased the *Chlorella* sp. cell concentration and electrolysis field strength to 0.43 ± 0.01 kg/m^3 and 13 V/cm, respectively. The total lipid extracted from C_{WOET} (8.18 ± 0.49% $g_{lipid}/g_{dry\ wt}$) was found to be 1.13-fold higher than C_{WET} (7.22 ± 0.47% $g_{lipid}/g_{dry\ wt}$). It was also observed that the final temperature of C_{WET} increased from 24.3 ± 0.1°C to 57.2 ± 0.1°C attributed by the increasing of current intensity. Meanwhile, when 0.46 ± 0.01 kg/m^3 *Chlorella* sp. was treated with ET method at 14 V/cm for 3000 s under a semi-continuous system and recycled at 2.3 $\mu m^3/s$. A similar trend of lipid was also attained which showed the total lipid of D_{WET} (7.72 ± 0.54 % $g_{lipid}/g_{dry\ wt}$) was 1.11-fold lower compared to D_{WOET} (8.58 ± 0.49% $g_{lipid}/g_{dry\ wt}$). The final temperature of D_{WET} was also increased from 24.4 ± 0.1°C to 59.1 ± 0.6°C which was also corresponded to the increasing of its current intensity during ET experiment.

The cause of low total lipid in C_{WET} and D_{WET} samples than the control sample might be because the treatment conditions used here had provided an extreme condition for the microalga cells. Further investigation regarding the experimental design are needed to clarify the results obtained here. For that reason, this extreme condition might contribute to the lipid oxidation phenomenon which may occur during the ET method (Daghrir et al., 2014). The formation of oxygen gases either from the surrounding or released from the anode terminal, may assist the lipid oxidation mechanism. According to Meullemiestre et al. (2016) the lipid degradation could also be affected by hydrolysis or oxidation mechanism such as photo-oxidation, auto-oxidation or enzymatic-oxidation reaction. It was also presumed that the increasing of the sample temperature might had assisted the lipid oxidation process. In other theory, during ET method the microalga cells wall was destroyed and led the intracellular lipid released to the media, then was removed during centrifugation, hence, lowering the total lipid extracted from these samples. As a matter

Table 1
ET operational parameters and Chlorella sp. total lipid content based on test conditions

Test conditions	Microalga conc. (kg/m^3)	Treatment time (s)	Field strength (V/cm)	Aeration flow ($\mu m^3/s$)	Recycling flowrate ($\mu m^3/s$)	Temperature (°C)		Total lipid (% $g_{lipid}/g_{dry\ wt}$)	Methyl palmitate (g/m^3)
						Initial	Final		
A_{WOET}	0.18 ± 0.01	-	-	-	-	22.5 ± 0.1	-	NA	NA
A_{WET}	0.18 ± 0.01	600	13	-	-	22.5 ± 0.1	29.0 ± 0.6	NA	NA
B_{WOET}	0.39 ± 0.02	-	-	-	-	24.3 ± 0.1	-	7.97 ± 0.43	NA
B_{WET}	0.39 ± 0.02	1800	5	16.7	-	24.3 ± 0.1	38.0 ± 0.9	7.95 ± 0.37	NA
C_{WOET}	0.43 ± 0.02	-	-	-	-	24.3 ± 0.1	-	8.18 ± 0.49	NA
C_{WET}	0.43 ± 0.02	1800	13	16.7	-	24.3 ± 0.1	57.2 ± 0.1	7.22 ± 0.47	NA
D_{WOET}	0.46 ± 0.01	-	-	-	-	24.4 ± 0.1	-	8.58 ± 0.49	0.049 ± 0.005
D_{WET}	0.46 ± 0.01	3000	14	-	2.3	24.4 ± 0.1	59.1 ± 0.6	7.72 ± 0.54	0.045 ± 0.005

Note: A_{WET} : Batch system without an air aeration flow and recycling flow; B_{WET} : Batch system with an air aeration flow but without recycling flow; C_{WET} : Batch system with an air aeration flow but without recycling flow; D_{WET} : Semi-continuous system without an air aeration flow but with recycling flow; NA: Not analysed

of fact, the trend of result obtained in this study was also similar to the results gained by Daghrir et al. (2014). They had also found that the total lipid of *C. vulgaris* WOET sample was 6.23 ± 1.05% $g_{lipid}/g_{dry\ wt}$ which was higher than WET sample about 5.53 ± 0.23% $g_{lipid}/g_{dry\ wt}$. They used an electrochemical treatment method which an adaptation of electrolysis method to extract lipid from *C. vulgaris*. Instead of using stainless steel, titanium (Ti) insulated with iridium (IV) oxide (Ti/IrO_2) was used as the anode material. The material promotes good interaction between the hydroxyl and the Ti/IrO_2 surface has a good mechanical strength compared to stainless steel (Zaviska et al., 2011). However, the market price of stainless steel is 1.4 times cheaper than Ti (Great American Group, 2015). Hence, this is one of the reasons stainless steel was used as the electrode material and may contribute to lower production costs when using the ET method.

Comparison of Total Lipid Extracted Using Electrical Based Approach

Table 2 shows the comparison of total lipid extracted from five microalgae species including the extractions data based on pervious reported literatures and the present study. Based on these data, the highest total lipid extracted that attained was 26.37 ± 0.47% $g_{lipid}/g_{dry\ wt}$ from *A. falcatus* which was 3.4 times higher than the total lipid obtained in this study (Misra et al., 2014). This might due to the higher microalga concentration (2.88 kg/m^3) that was used. Moreover, the cell size of *A. falcatus* was much bigger than *Chlorella* sp. cell. In fact, the electrode type (i.e. non-sacrificial carbons) used was also different and noted that the conductivity of carbon (5.9×10^6 S/m) was 4 times higher compared to stainless steel (1.37×10^6 S/m) (TIBTECH). By all means, higher conductivity would provides a good efficiency of the cell membrane electroporation (Joannes et al., 2015b).

Hua et al. (2016) used an electrochemical advanced oxidation processes to extract lipid from *S. dimorphus*. The microalga was treated at 8 V/cm and 0.5 A for 7200 s and using stainless steel (cathode) and Ti_4O_7 (anode). The results indicated that the total lipid extracted was increased from 15.2 ± 0.6% to 23.4 ± 0.7% $g_{lipid}/g_{dry\ wt}$. Using the non-sacrificial carbons as the electrodes and applying 1.0 V/cm and 1 A for 1800 s to *A. falcatus*, the highest total lipid can be obtained was 26.37 ± 0.47% $g_{lipid}/g_{dry\ wt}$ (Guldhe et al., 2016). Meanwhile, exposing *S. obliquus* at 1.7 V/cm and 1.5 A for 3600 mins, the highest total lipid achieved was 15.42 ± 0.19% $g_{lipid}/g_{dry\ wt}$. However, decreasing the field strength at 0.77 V/cm and adding 6 kg/m^3 of NaCl (other parameters remained the same), the total lipid of *S. obliquus* extracted was increased to 16.2 ± 0.27% $g_{lipid}/g_{dry\ wt}$ (Misra et al., 2015, 2014). Misra et al. (2014) also found that *C. sorokiniana* lipid could be extracted up to 13.72 ± 0.43% $g_{lipid}/g_{dry\ wt}$ at 0.6 V/cm and 1.0 A for 3600s. Daghrir et al. (2014) had also added 1.42 kg/m^3 of Na_2SO_4 and treated *C. vulgaris* at 14.3 V/cm, 0.6 A and 6000s. In addition, the microalga was recycled at 6.6 $\mu m^3/s$. By using stainless steel at cathode and Ti/IrO_2 at anode, the total lipid extracted was around 5.53 ± 0.23% $g_{lipid}/g_{dry\ wt}$.

Table 2
Comparison of proposed ET method performance with other literatures

Microalga	Parameters	Optimum values	Reference
Chlorella sp.	Initial total lipid	8.58 ± 0.49% $g_{lipid}/g_{dry\ wt}$	Present study
	Final total lipid	7.72 ± 0.54% $g_{lipid}/g_{dry\ wt}$	
	Field strength	14 V/cm	
	Current intensity	1.92 ± 0.02 A	
	Treatment time	3000 s	
	Recycling flow	2.3 $\mu m^3/s$	
	Microalga conc.	0.46 ± 0.01 kg/m^3	
	Electrode type	Stainless steel (cathode and anode)	
	Solvent extraction	Folch method	
Scenedesmus dimorphus	Initial total lipid	15.2 ± 0.6% $g_{lipid}/g_{dry\ wt}$	Hua et al. (2016)
	Final total lipid	23.4 ± 0.7% $g_{lipid}/g_{dry\ wt}$	
	Field strength	8 V/cm	
	Current intensity	0.5 A	
	Treatment time	7200 s	
	Microalga conc.	1.4 kg/m^3	
	Electrode type	Stainless steel (cathode), $Ti4O_7$ (anode)	
	Solvent extraction	Dichloromethane and methanol	
Ankistrodesmus falcatus	Initial total lipid	Not stated	Guldhe et al. (2016)
	Final total lipid	26.37 ± 0.47% $g_{lipid}/g_{dry\ wt}$	
	Field strength	1.0 V/cm	
	Current intensity	1.0 A	
	Treatment time	1800 s	
	Microalga conc.	2.88 kg/m^3	
	Electrode type	non-sacrificial carbons (cathode and anode)	
	Solvent extraction	Folch method	
Scenedesmus obliquus	Initial total lipid	13.60 ± 0.16% $g_{lipid}/g_{dry\ wt}$	Misra et al. (2015)
	Final total lipid	16.20 ± 0.27% $g_{lipid}/g_{dry\ wt}$	
	Field strength	0.77 V/cm	
	Current intensity	1.5 A	
	Treatment time	3600 s	
	Microalga conc.	2.88 kg/m^3	
	Electrode type	Non-sacrificial carbons (cathode and anode)	
	Solvent extraction	Folch method	
	Electrolyte	6 kg/m^3 of NaCl	
Chlorella sorokiniana	Initial total lipid	12.37 ± 0.17% $g_{lipid}/g_{dry\ wt}$	Misra et al. (2014)
	Final total lipid	13.72 ± 0.43% $g_{lipid}/g_{dry\ wt}$	
	Field strength	0.6 V/cm	
	Current intensity	1.0 A	
	Treatment time	3600 s	
	Microalga conc.	2.8 kg/m^3	
	Electrode type	Non-sacrificial carbons (cathode and anode)	
	Solvent extraction	Folch method	
	Electrolyte	6 kg/m^3 of NaCl	

Table 2 *(continue)*

Microalga	Parameters	Optimum values	Reference
Scenedesmus obliquus	Initial total lipid	15.56 ± 0.06% $g_{lipid}/g_{dry\ wt}$	Misra et al. (2014)
	Final total lipid	15.42 ± 0.19% $g_{lipid}/g_{dry\ wt}$	
	Field strength	1.7 V/cm	
	Current intensity	1.5 A	
	Treatment time	3600 s	
	Microalga conc.	2.8 kg/m^3	
	Electrode type	Non-sacrificial carbons (cathode and anode)	
	Solvent extraction	Folch method	
Chlorella vulgaris	Initial total lipid	6.23 ± 1.05% $g_{lipid}/g_{dry\ wt}$	Daghrir et al. (2014)
	Final total lipid	5.53 ± 0.23% $g_{lipid}/g_{dry\ wt}$	
	Field strength	14.3 V/cm	
	Current intensity	0.6 A	
	Treatment time	3000 s	
	Recycling flow	6.6 µm^3/s	
	Microalga conc.	Not stated	
	Electrode type	Stainless steel (cathode), Ti/IrO_2 (anode)	
	Solvent extraction	Folch method	
	Electrolyte	1.42 kg/m^3 of Na_2SO_4	

From this Table of comparison, different treatments with different microalga species will provide different results. However, from Table 2 it was also observed that the total lipid of untreated microalga was always higher or slightly higher than the total lipid of treated microalga. Hence, further investigation should be implemented (e.g. experimental design parameters and optimization) in order to make the ET method to become feasible to be used in future microalga lipid extraction for biodiesel production.

Effect of Direct Heat Treatment

In order to determine the effect of heat towards the microalga lipid extraction, a simple experiment was conducted. Table 3 shows the amount of total lipid extracted from *Chlorella* sp. with different temperature of treatment. The total lipid extracted at 30°C, 45°C and 60°C were 8.59 ± 0.46% $g_{lipid}/g_{dry\ wt}$, 8.57 ± 0.58% $g_{lipid}/g_{dry\ wt}$ and 8.55 ± 0.57% $g_{lipid}/g_{dry\ wt}$, respectively. It can be concluded that, the total lipid extracted in this experiments were no significant difference. According to Zbinden et al. (2013) microalga cell could not be lysed using a direct treatment at 50°C. However, exposing the microalga cell to higher temperature as high as 70°C could cause the cell to lyse (Madigan et al., 2009). In this study, increasing the temperature up to 60°C would not lyse the *Chlorella* sp. cell wall as the total lipid extracted was approximately the same as the control sample (D_{WOET}). It can be concluded that, direct heating will not cause the disruption of microalga cell but may weaken its cell wall.

Table 3
Effect of direct heat treatment on Chlorella sp. lipid extraction

Test condition	Microalga conc. (kg/m^3)	Temperature (^{o}C)	Total lipid (% $g_{lipid}/g_{dry\ wt}$)
D_{WOET}	0.46 ± 0.01	24	8.58 ± 0.49
H_1	0.46 ± 0.01	30	8.59 ± 0.46
H_2	0.46 ± 0.01	45	8.57 ± 0.58
H_3	0.46 ± 0.01	60	8.55 ± 0.57

Microalga Lipid Functional Groups

The FTIR results were shown in Figure 2. From Figure 2(a) it was observed that a similar FTIR spectrum was obtained from D_{WOET} and D_{WET}. Only water functional group was presented in both samples. This had also indicated that lipid was not spontaneously released to the surrounding (Goettel et al., 2013). However, after performing the solvent extraction method, it was found that the results had confirmed the existence of lipid functional groups from both samples as shown in Figure 2(b). The lipid functional groups detected were C = O bond (1740 cm^{-1}), C – O bond (1380 cm^{-1}) and C – O – C bond (1250–1070 cm^{-1}) in *Chlorella* sp. from D_{WOET} and D_{WET} samples (Dean et al., 2010; Forfang et al., 2017). From these results, it can be stated that the microalga lipid cannot be extract directly if using the ET method alone. Thus, it required solvents to access into the lipid granule and then extracting the lipid from the microalga cells.

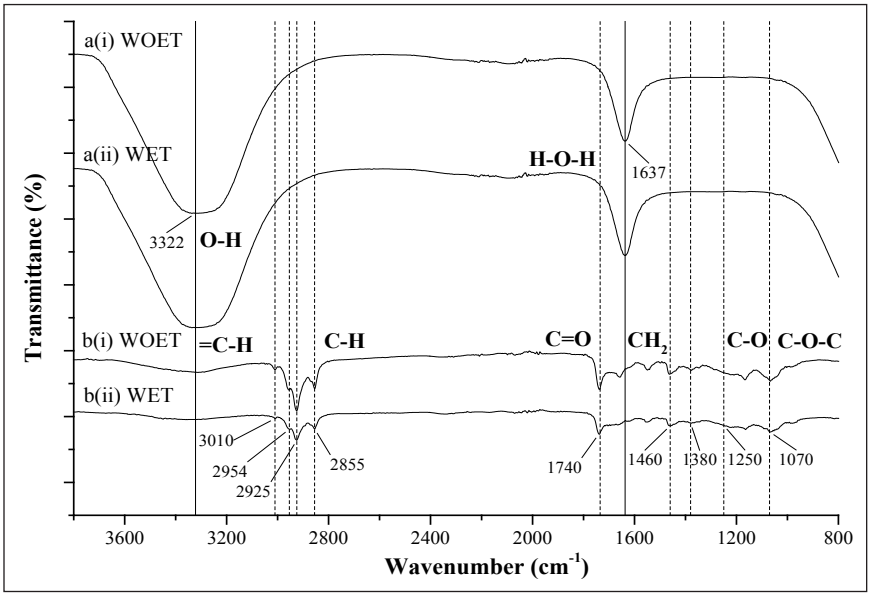


Figure 2. (a) FTIR spectra of (i) *Chlorella* sp. D_{WOET} and (ii) *Chlorella* sp. D_{WET} prior to solvent extraction. (b) FTIR spectra from (i) *Chlorella* sp. D_{WOET} and (ii) *Chlorella* sp. D_{WET} after solvent extraction

Surface Morphology

Figure 3 shows the surface morphology of *Chlorella* sp. cells from D_{WOET} and D_{WET} samples. The cells in Figure 3(a) and 3(b) clearly demonstrated a distinct cell envelope whereas the cells in Figure 3(c) to 3(f) show cell fragments and were agglomerated. The result obtained here was different from the result reported by Yap and co-workers that have observed unshrunk *Chlorella* sp. cells with spherical shape (Yap et al., 2014). The microalga cells in Figure 3(a) and 3(b) appeared to be shrunk and shrivelled due to the freeze drying effect

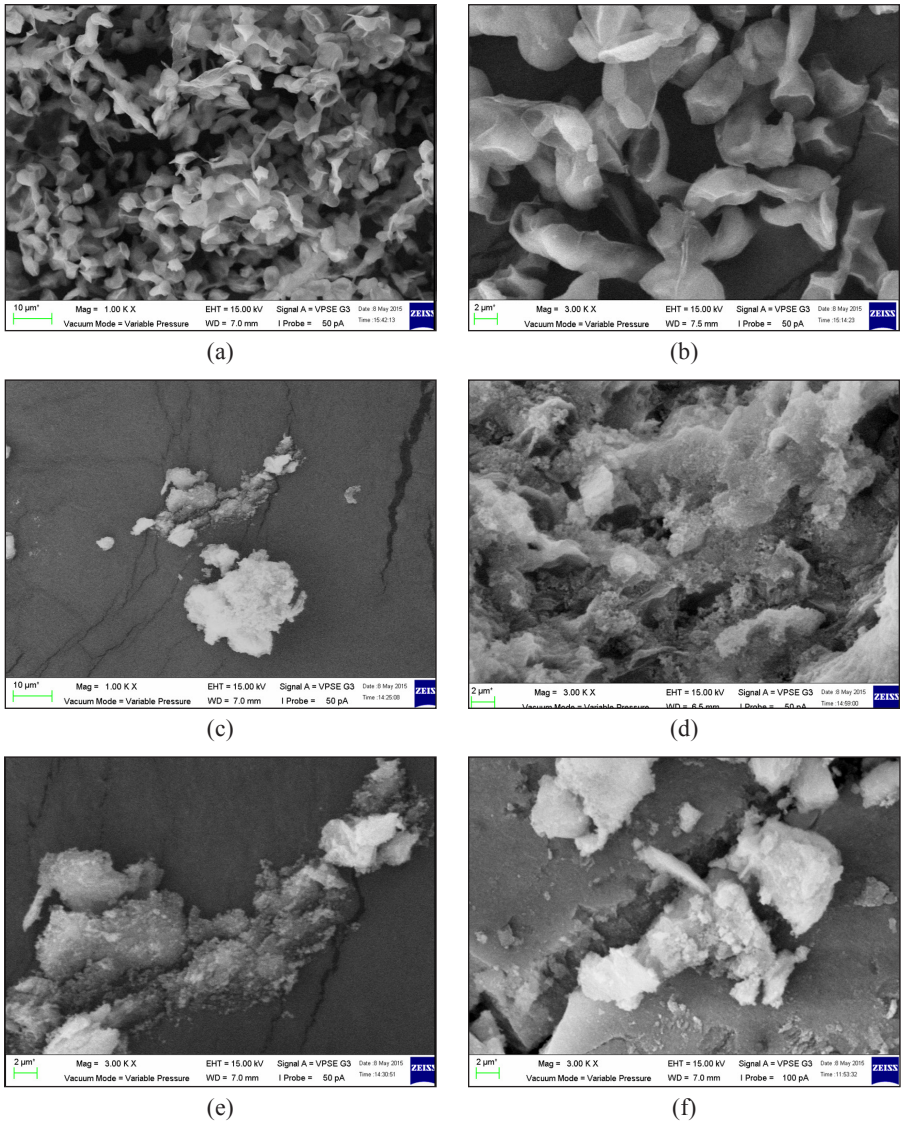


Figure 3. Freeze-dried *Chlorella* sp. cells from D_{WOET} sample viewed under a SEM: (a) at 1000×; and (b) at 3000× magnification power. Freeze-dried *Chlorella* sp. cells D_{WET} sample viewed under a SEM: (c) at 1000×; and (d, e, f) at 3000× magnification power.

(Joannes et al., 2016). Although the microalga cells do not have their typical round shape, but the cell membranes appear unbroken and intact. Meanwhile, Figure 3(c) to 3(f) show only cell fragments with no remaining intact cells. It can be concluded that the cells were disrupted by the external potential applied during ET method.

It is important to be able to view the before and after the condition of the cell walls to claim any enlargement of cell membrane pores. Although it was most likely that the treatment completely destroyed the cell membrane over 3000 s, freeze drying could also have had an added effect on the treated cells. Noted that, freeze drying method also contributes the weakening of the cell membrane but the process will not destroy the cells (Lee et al., 2012). Even though, the *Chlorella* sp. cells were fully destroyed by ET method, their lipid was not spontaneously released. Therefore, a solvent was required to extract the lipid remaining inside the cell in a fashion similar to electrochemical treatment and PEF work (Daghrir et al., 2014; Eing et al., 2013).

FAME Compositions

The FAME compounds of *Chlorella* sp. from D_{WOET} and D_{WET} samples were identified using GCMS and predicted based on the National Institute of Standards and Technology (NIST) library database. Table 3 shows the FAME of the *Chlorella* sp. in D_{WET} sample. Based on the GCMS analysis, three significant peaks were obtained at retention times of 15.90, 16.98 and 22.46 as presented in Figure 4. The GCMS profile for D_{WOET} was not reported here due to the similar mass spectrum obtained thus, only the GCMS profile of D_{WET} was presented.

The GCMS profile confirmed the existence of two major components found at retention times of 16.98 and 22.46 as shown in Figure 5(a) and Figure 5(b), respectively. The two

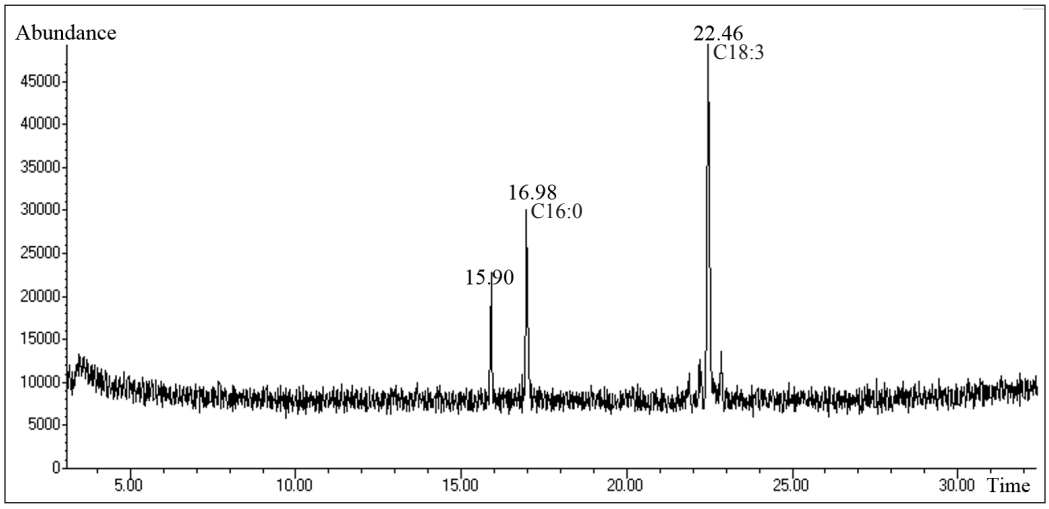
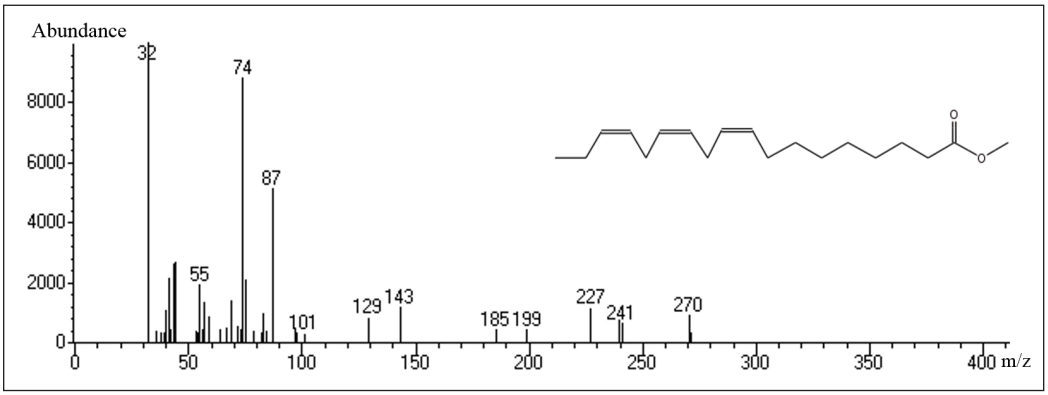
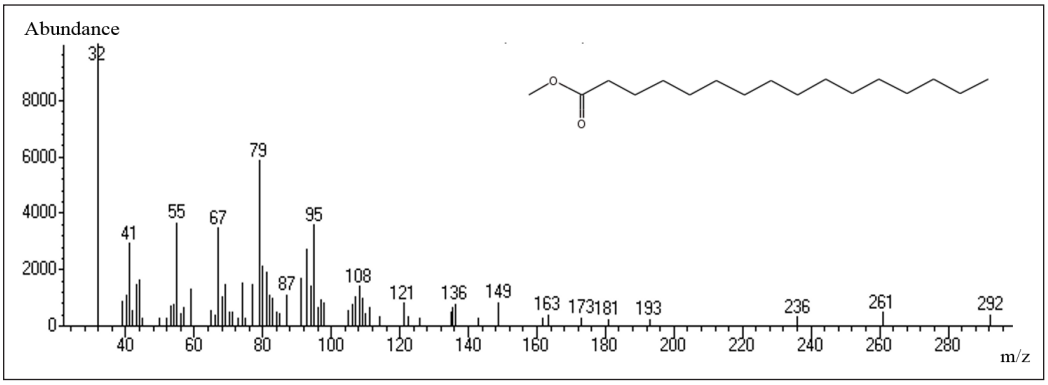


Figure 4. GCMS chromatogram of FAME from *Chlorella* sp. D_{WET}

FAME were methyl palmitate (Hexadecanoic acid methyl ester) and methyl linolenate (Z, Z, Z-9, 12, 15-Octadecatrienoic acid methyl ester). Methyl palmitate is a monosaturated FAME and is mostly present in biodiesel composition, whereas methyl linolenate is a polyunsaturated FAME which has three double bonds unfavourable for biodiesel properties.



(a)



(b)

Figure 5. Mass spectrum of (a) methyl palmitate at retention time 16.98 and (b) methyl linolenate at retention time 22.46 from *Chlorella* sp. D_{WET}

Table 4
FAME components analysed from Chlorella sp. D_{WET} sample

FAME Name	Hexadecanoic acid methyl ester	(Z, Z ,Z)-9,12,15-Octadecatrienoic acid methyl ester
Common Name	Methyl palmitate (C16:0)	Methyl linolenate (C18:3)
Molecular Formula	$C_{17}H_{34}O_2$	$C_{19}H_{32}O_2$
Molecular Weight	270	292
Peak Area (%)	25.97	60.19

From Table 4, it can be seen that methyl linolenate (60.19%) was a dominant component than methyl palmitate (25.97%). Higher levels of polyunsaturated fatty acids will contribute to instability of oxidation in biodiesel engines (Mahapatra et al., 2013). It is important to note that monosaturated fatty acids are preferred for use in biodiesel engines due to their ability to provide a high cetane number for excellent ignition and oxidation stability (Naik et al., 2010; Patel et al., 2017; Rasoul-Amini et al., 2011). It was also observed that this local *Chlorella* sp. lipid was very similar in composition to vegetable oils such as rapeseed (mainly consisting of C16 and C18) and had comparable biodiesel properties (Lang et al., 2001).

The concentration of methyl palmitate extracted from C*hlorella* sp. in D_{WOET} and D_{WET} samples was presented in Table 1. Based on the standard curve of methyl palmitate, the amount of methyl palmitate extracted from the *Chlorella* sp. D_{WOET} and D_{WET} samples were 0.049 ± 0.005 g/m^3 and 0.045 ± 0.005 g/m^3, respectively. In term of number, the amount of methyl palmitate attained *Chlorella* sp. D_{WOET} was slightly higher than in *Chlorella* sp. D_{WET} can be corresponded to the higher total lipid was also attained from the sample (Mansa et al., 2018). However, these values showed that there was no significant difference in methyl palmitate concentration in both samples. The upper layer of the *Chlorella* sp. in D_{WET} sample was also tested for lipid determination using the gravimetric method, transesterified and analysed for FAME. Unfortunately, it was found that the upper layer contained insignificant amounts of lipid and FAME, indicating that *Chlorella* sp. lipid droplets were not spontaneously released into the upper surface and remained inside the cells (Eing et al., 2013).

CONCLUSION

In this work, a local Borneo microalga *Chlorella* sp. was chosen as a potential biodiesel feedstock and ET method as cell disruption method prior to solvent extraction. Using ET method described here, the highest total lipid extracted from 0.46 ± 0.01 kg/m^3 *Chlorella* sp. was $7.72 \pm 0.54\%$ $g_{lipid}/g_{dry\ wt}$ with treatment conditions using the field strength of 14 V/cm and recycled at 2.3 µm^3/s for 3000 s. Heating the microalga to 60^{o}C will not destroy the *Chlorella* sp. cell wall. Indicating that the destroyed and agglomerated cells in D_{WET} sample was attributed by the external potential which was the electric fields in ET. The GCMS analysis also revealed that methyl linolenate (C18:3) was dominant compound than methyl palmitate (C16:0). Methyl palmitate was presented at a level of 0.045 ± 0.005 g/m^3 in $D_{WET.}$ *Chlorella* sp. D_{WET} from the upper layer contained insignificant amounts of lipid and FAME, explaining why the lipid droplets were not spontaneously released from the microalga cells. Hence, a solvent extraction method was required to extract the remaining lipid inside the cell wall. The lipid layer was expected to be formed on the upper

layer of the treated sample. However, based on visual inspection and FTIR analysis, this lipid layer was not present. Based on the results attained in this study, it was found that ET method was unable to enhance the lipid extraction of *Chlorella* sp. Even though the expected result was not achieved, there are still lots of improvements that can be done for the ET to become feasible to be used as the pre-treatment method in microalga lipid extraction. For future works, the ET system with or without air aeration and recycling flow should be emphasized. This is to determine which system gives the best result in term of less temperature increment and with high total lipid extraction. Apart from that, the ET parameters such as treatment time, field strength, recycling flowrate and solvent type need to be varied in order to investigate the optimum point of the microalga lipid extraction. The effect of lipid oxidation needs further investigation to understand the correlation between the extracted lipid and loss of lipid.

ACKNOWLEDGMENTS

The authors wish to acknowledge the Ministry of Agriculture Malaysia a Seaweed Research Unit Grant GPRL 017 (Biochemical Process of Seaweed for Industrial Product) and Universiti Malaysia Sabah (Grant No: SBK0235-TK-2015) for the financial support of the study.

REFERENCES

Abdullah, N., Amran, N. A., & Yasin, N. H. M. (2017). Algae oil extraction from freshwater microalgae *Chlorella vulgaris*. *Malaysian Journal of Analytical Sciences*, *21*(3), 735–744.

Barros, A. I., Goncalves, A. L., Simoes, M., & Pires, J. C. M. (2015). Harvesting techniques applied to microalgae: A review. *Renewable and Sustainable Energy Reviews*, *41*, 1489–1500.

BP Statistical Review of World Energy, E. (2017). *BP Statistical Review of World Energy June 2017*. Retrieved June 3, 2017, from bp.com/statisticalreview

Chen, L., Li, R., Ren, X., & Liu, T. (2016). Improved aqueous extraction of microalgal lipid by combined enzymatic and thermal lysis from wet biomass of *Nannochloropsis oceanica*. *Bioresource Technology*, *214*, 138–143.

Daghrir, R., Igounet, L., Brar, S. K., & Drogui., P. (2014). Novel electrochemical method for the recovery of lipids from microalgae for biodiesel production. *Journal of the Taiwan Institute of Chemical Engineers*, *45*(1), 153–162.

Dean, A. P., Sigee, D. C., Estrada, B., & Pittman, J. K. (2010). Using FTIR spectroscopy for rapid determination of lipid accumulation in response to nitrogen limitation in freshwater microalgae. *Bioresource Technology*, *101*(12), 4499–4507.

Dimitrov, D. S. (1995). *Electroporation and electrofusion of membranes*. In Lipowsky, R., & Sackmann. E. (Eds.), *Handbook of Biological Physics* (Vol. 1, pp. 851–898). Amsterdam, Elsevier Science.

Eing, C., Goettel, M., Straessner, R., Gusbeth, C., & Frey, W. (2013). Pulsed electric field treatment of microalgae - benefits for microalgae biomass processing. *IEEE Transactions on Plasma Science*, *41*(10), 2901–2907.

Folch, J., Lees, M., & Sloane-Stanley, G. H. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *Journal of Biological Chemistry, 226*(1), 497–509.

Forfang, K., Zimmermann, B., Kosa, G., Kohler, A., & Shapaval, V. (2017). FTIR spectroscopy for evaluation and monitoring of lipid extraction efficiency for oleaginous fungi. *Plos One*, *12*(1), 1–17.

Goettel, M., Eing, C., Gusbeth, C., Straessner, R., & Frey, W. (2013). Pulsed electric field assisted extraction of intracellular valuables from microalgae. *Algal Research*, *2*(4), 401–408.

Great American Group. (2015). *Metals Monitor.* Retrieved June 3, 2017, from http://www.greatamerican.com.

Grimi, N., Dubois, A., Marchal, L., Jubeau, S., Lebovka, N. I., & Vorobiev, E. (2014). Selective extraction from microalgae *Nannochloropsis* sp . using different methods of cell disruption. *Bioresource Technology*, *153*, 254–259.

Guldhe, A., Misra, R., Singh, P., Rawat, I., & Bux, F. (2016). An innovative electrochemical process to alleviate the challenges for harvesting of small size microalgae by using non-sacrificial carbon electrodes. *Algal Research*, *19*, 292–298.

Günerken, E., D'Hondt, E., Eppink, M. H. M., Garcia-Gonzalez, L., Elst, K., & Wijffels, R. H. (2015). Cell disruption for microalgae biorefineries. *Biotechnology Advances*, *33*(2), 243–260.

Hua, L., Guo, L., Thakkar, M., Wei, D., Agbakpe, M., Kuang, L., ... & Zhang, W. (2016). Effects of anodic oxidation of a substoichiometric titanium dioxide reactive electrochemical membrane on algal cell destabilization and lipid extraction. *Bioresource Technology*, *203*, 112–117.

Japar, A. S., Takriff, M. S., & Mohd Yasin, N. H. (2017). Harvesting microalgal biomass and lipid extraction for potential biofuel production : A review. *Journal of Environmental Chemical Engineering*, *5*(1), 555–563.

Jeevan-Kumar, S. P., Garapati, V. K., Dash, A., Scholz, P., & Banerjee, R. (2017). Sustainable green solvents and techniques for lipid extraction from microalgae : A review. *Algal Research*, *21*, 138–147.

Joannes, C., Mansa, R. F., Sipaut, C. S., Yasir, S. M., & Dayou, J. (2016). Comparative studies of cell growth of freshwater microalga *Chlorella* sp. in photoautotrophic, heterotrophic and mixotrophic cultures. *Jurnal Teknologi (Science & Engineering)*, *78*(7), 83–89.

Joannes, C., Sipaut, C. S., Dayou, J., Yasir, S. M., & Mansa, R. F. (2015a). Review paper on cell membrane electroporation of microalgae using electric field treatment method for microalgae lipid extraction. In *IOP Conference Series: Materials Science and Engineering* (Vol. 78, No. 1, p. 012034). Miri, Sarawak, Malaysia: IOP Publishing.

Joannes, C., Sipaut, C. S., Dayou, J., Yasir, S. M., & Mansa, R. F. (2015b). The potential of using pulsed electric field (PEF) technology as the cell disruption method to extract lipid from microalgae for biodiesel production. *International Journal of Renewable Energy Research*, *5*(2), 598–621.

Kotnik, T., Kramar, P., Pucihar, G., Miklavčič, D., & Tarek, M. (2012). Cell membrane electroporation - Part 1: The phenomenon. *IEEE Electrical Insulation Magazine*, *28*(5), 14–23.

Kumar, M., & Sharma, M. P. (2016). Selection of potential oils for biodiesel production. *Renewable and Sustainable Energy Reviews*, *56*, 1129–1138.

Lam, G. P., Postma, P. R., Fernandes, D. A., Timmermans, R. A. H., Vermuë, M. H., Barbosa, M. J., ... & Olivieri, G. (2017). Pulsed electric field for protein release of the microalgae *Chlorella vulgaris* and *Neochloris oleoabundans*. *Algal Research*, *24*, 181–187.

Lee, A. K., Lewis, D. M., & Ashman, P. J. (2012). Disruption of microalgae cells for the extraction lipids for biofuels: Process and specific energy requirements. *Biomass and Bioenergy*, *46*, 89–101.

Madigan, M., Martinko, J., Dunlap, P., & Clark, D. (2009). *Brock biology of microorganisms*. San Francisco: Pearson Benjamin Cummings.

Mahapatra, D. M., Chanakya, H. N., & Ramachandra, T. V. (2013). *Euglen*a sp. as a suitable source of lipids for potential use as biofuel and sustainable wastewater treatment. *Journal of Applied Phycology*, *25*(3), 855–865.

Mansa, R. F., Sipaut, C. S., Yasir, S., Dayou, J. & Joannes, C. (2018). Comparative Studies of Cell Growth, Total Lipid and Methyl Palmitate of Ankistrodesmus sp. In Phototrophic, Mixotrophic and Heterotrophic Cultures For Biodiesel Production. *International Journal of Renewable Energy Research*, *8*(1), 438–450.

Mansa, R. F., Tahir, A., Hua, L. M., Dayou, J., & Sipaut, C. S. (2012). Design of a pilot scale outdoor photobioreactor for mass cultivation of local microalga. *International Journal of Engineering and Physical Sciences*, *6*, 348–352.

Meullemiestre, A., Breil, C., Abert-Vian, M., & Chemat, F. (2016). Microwave, ultrasound, thermal treatments, and bead milling as intensification techniques for extraction of lipids from oleaginous *Yarrowia lipolytica* yeast for a biojetfuel application. *Bioresource Technology*, *211*, 190–199.

Misra, R., Guldhe, A., Singh, P., Rawat, I., & Bux, F. (2014). Electrochemical harvesting process for microalgae by using nonsacrificial carbon electrode: a sustainable approach for biodiesel production. *Chemical Engineering Journal*, 255, 327–333.

Misra, R., Guldhe, A., Singh, P., Rawat, I., Stenström, T. A., & Bux, F. (2015). Evaluation of operating conditions for sustainable harvesting of microalgal biomass applying electrochemical method using non sacrificial electrodes. *Bioresource Technology*, *176*, 1–7.

Naik, S. N., Goud, V. V., Rout, P. K., & Dalai, A. K. (2010). Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, *14*(2), 578–597.

Parniakov, O., Barba, F. J., Grimi, N., Marchal, L., Jubeau, S., Lebovka, N., & Vorobiev, E. (2015). Pulsed electric field assisted extraction of nutritionally valuable compounds from microalgae *Nannochloropsis* spp. using the binary mixture of organic solvents and water. *Algal Research*, 8, 128–134.

Patel, A., Gami, B., Patel, P., & Patel, B. (2017). Microalgae : Antiquity to era of integrated technology. *Renewable and Sustainable Energy Reviews*, *71*, 535–547.

Postma, P. R., Pataro, G., Capitoli, M., Barbosa, M. J., Wijffels, R. H., Eppink, M. H. M., ... & Ferrari, G. (2016). Selective extraction of intracellular components from the microalga *Chlorella vulgaris* by combined pulsed electric field-temperature treatment. *Bioresource Technology*, *203*, 80–88.

Rasoul-Amini, S., Montazeri-Najafabady, N., Mobasher, M. A., Hoseini-Alhashemi, S., & Ghasemi, Y. (2011). *Chlorella* sp.: A new strain with highly saturated fatty acids for biodiesel production in bubble-column photobioreactor. *Applied Energy*, *88*(10), 3354–3356.

Suali, E., & Sarbatly, R. (2012). Conversion of microalgae to biofuel. *Renewable and Sustainable Energy Reviews*, *16*, 4316–4342.

TIBTECH. Menu Metal Properties. Retrieved June 6, 2017, from http://www.tibtech.com/conductivity.php

Uduman, N., Qi, Y., Danquah, M. K., Forde, G. M., & Hoadley, A. (2010). Dewatering of microalgal cultures: A major bottleneck to algae-based fuels. *Journal of Renewable and Sustainable Energy*, *2*(1), 1–15.

Velasquez-Orta, S. B., Lee, J. G. M., & Harvey, A. P. (2013). Evaluation of fame production from wet marine and freshwater microalgae by in situ transesterification. *Biochemical Engineering Journal*, *76*, 83–89.

Yap, B. H. J., Crawford, S. A., Dumsday, G. J., Scales, P. J., & Martin, G. J. O. (2014). A mechanistic study of algal cell disruption and its effect on lipid recovery by solvent extraction. *Algal Research*, *5*(1), 112–120.

Zaviska, F., Drogui, P., Blais, J. F., Mercier, G., & Lafrance, P. (2011). Experimental design methodology applied to electrochemical oxidation of the herbicide atrazine using Ti/IrO_2 And Ti/SnO_2 circular anode electrodes. *Journal of Hazardous Materials*, *185*, 1499–1507.

Zbinden, M. D. A., Sturm, B. S. M., Nord, R. D., Carey, W. J., Moore, D., Shinogle, H., & Stagg-Williams, S. M. (2013). Pulsed electric field (PEF) as an intensification pretreatment for greener solvent lipid extraction from microalgae. *Biotechnology and Bioengineering*, *110*(6), 1605–1615.

Zhou, W., Gao, L., Cheng, W., Chen, L., Wang, J., Wang, H., ... & Liu, T. (2016). Electro-flotation of *Chlorella* sp. assisted with flocculation by chitosan. *Algal Research*, *18*, 7–14.

Pertanika J. Sci. & Technol. 26 (4): 1713 - 1724 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Routing Protocol Optimization of Smart Grid Neighborhood Area Network based on IEEE 802.11 Standard

Chhaya, L. K. [1]*, Sharma, P. [2], Kumar, A. [1] and Bhagwatikar, G. [3]

[1]Department of Electronics, Instrumentation and Control Engineering, University of Petroleum and Energy Studies, Dehradun 248007, India
[2]Pandit Deendayal Petroleum University, Gujarat 382421, India
[3]SANY Group, B-604, Kumar Peninsula, Baner Pashan Link Road, Pashan,Pune, Maharashtra, 411021 , India

ABSTRACT

A power 'Grid' is a network that carries electricity from power plants to consumer premises. The grid is made 'smart' as it can monitor and control the distribution system by taking intelligent decisions. Smart Grid is an automated and broadly distributed energy generation, transmission and distribution network. Smart Grid network integrates an electrical distribution system with information and communication network. Communication network protocols are engineered, developed and established based on the layered approach. Each layer is designed to serve a specific functionality in collaboration with other layers. Layered approach for wired communication approach can be modified with cross layer approach for wireless communication for performance enhancement. Smart grid technology comprises of hierarchical and heterogeneous network with diverse set of communication protocols. This demands a divergence from primitive approach and adaptation of an innovative approach. This paper describes network design and optimization of routing protocol for Smart grid Neighborhood Area Network using Riverbed-OPNET software. A Cross layer approach is considered in parameter optimization of IEEE 802.11 standard. The proposed work shows parameter optimization of routing protocol for better network performance using simulation approach.

ARTICLE INFO

Article history:
Received: 11 December 2017
Accepted: 31 May 2018
Published: 24 October 2018

E-mail addresses:
lipi.chhaya@gmail.com (Chhaya, L. K.)
paawan.sharma@gmail.com (Sharma, P.)
ADESHKUMAR@DDN.UPES.AC.IN (Kumar, A.)
gowind.india@gmail.com (Bhagwatikar, G.)
* Corresponding author

Keywords: Communication, cross layer optimization, home area network, IEEE 802.11, neighborhood area network, smart grid, wide area network

INTRODUCTION

Smart grid is a multi-faceted technology with integration of electrical and communication infrastructure. Smart grid communication

ISSN: 0128-7680
e-ISSN: 2231-8526

network is a hierarchical and heterogeneous network containing various network layers. Data acquisition for monitoring and control requires diverse set of communication standards based on various applications. Wired as well as wireless communication standards can be used for Smart grid communications (Gungor et al., 2010). Smart grid communication architecture can be designed and optimized on the basis of three network layers based on bandwidth requirement, coverage area and application. These three network layers can be classified as Home Area Network (HAN), Neighborhood Area Network (NAN) and Wide Area Network (WAN). These layers are interconnected for real time monitoring, control, diagnosis and management of entire grid (Saputro et al., 2012). Each network requires an optimization of specific set of communication standards. Layered architecture supports various applications such as sensing, measurement, smart metering, home automation, SCADA and IoT etc. HAN covers home area monitoring, regulation, control and management. NAN is applicable for distributed generation and distribution automation. WAN shelters HAN and NAN for monitoring and control of entire communication network (Farooq & Jung, 2014). WAN is a gigantic network covering management of generation, transmission, distribution and utilization of entire grid (Mahmood et al., 2015). Smart grid data communication is characterized by intra-network and inter-network communications for operation and management of various components of grid (Erol-Kantarci & Mouftah, 2011). Figure 1 shows the layered communication infrastructure of Smart grid.

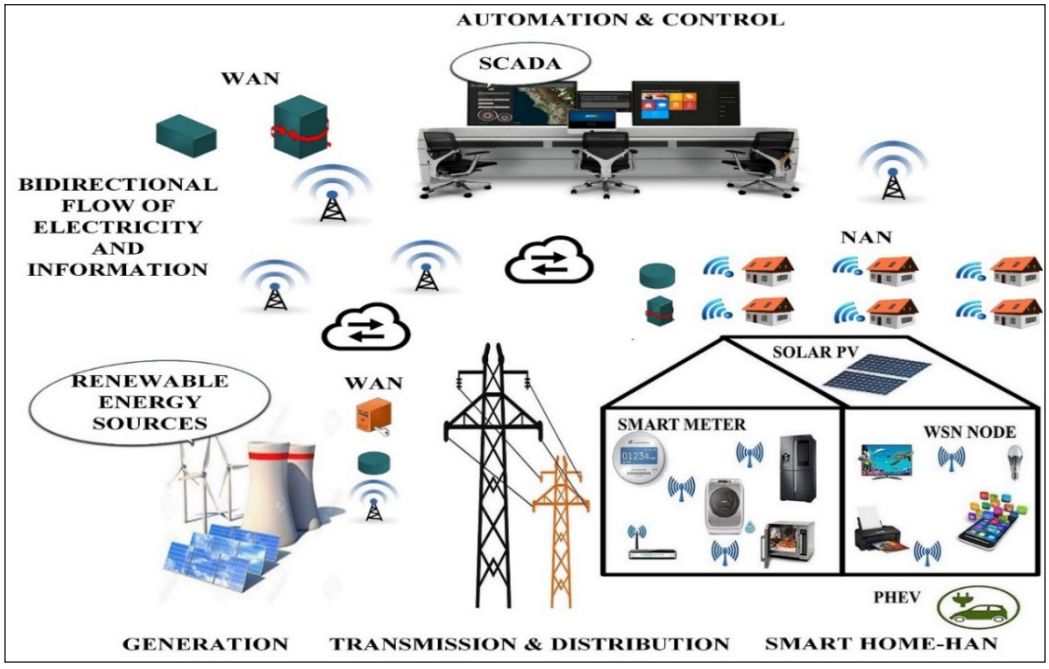


Figure 1. Layered Smart grid infrastructure

Each layer of Smart grid network is optimized for different set of communication protocols (Chen et al., 2006). The various choices available are IEEE 802.11, IEEE 802.15.1, IEEE 802.15.4 and IEEE 802.16 (Sun et al., 2009). Various communication protocols can be used for monitoring and control of various networks (Xu et al., 2014). Communication protocols are chosen for a specific application on the basis of various factors such as data rate, coverage area, and frequency spectrum (Patel et al., 2011).

Application of IEEE 802.11 Standard in Smart Grid

IEEE 802.11 standard is apt for Smart grid applications due to its features such as extensive availability around the world, plug and play devices, high data rates up to 600 Mbps, low cost, operability with IoT applications without protocol translation etc. Security of WLAN network is a crucial issue. IEEE 802.11 uses mainly three security protocols such as WEP (Wired Equivalent Privacy), WPA (Wi-Fi Protected Access) and WPA2 (Wi-Fi Protected Access, version-2) as shown in Figure 2. Various features of these security protocols are depicted in Table 1.

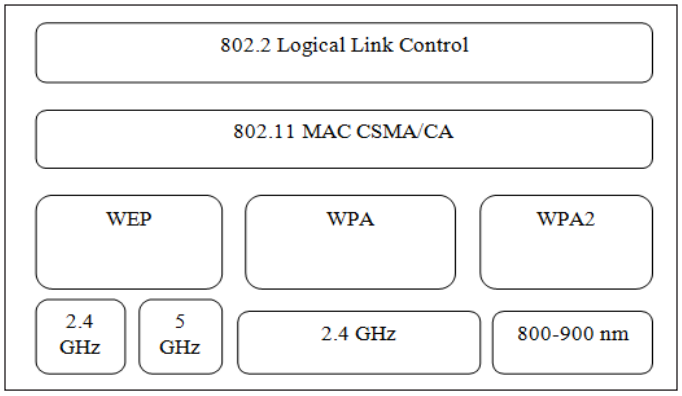


Figure 2. WLAN security protocols

Table 1
Comparison of various WLAN security protocols

PARAMETER	WEP	WPA	WPA2
Key Length	40/104 bits	128 bits	128 bits or higher
Encryption	RC4	TKIP	AES
Key management	Lack of Key Management	Robust mechanism, 4 way handshake	Robust mechanism, 4 way handshake
Data Integrity	CRC-32	Message Integrity code is generated	Cipher Block Chaining Message Authentication Code
Authentication	WEP-Open and Shared	WPA-PSK	WPA-2-Personal & Enterprise
Protection against Replay Attacks	No provision for protection	Sequence counter is implemented	48 bit packet number for attack prevention

From above table, it is apparent that WPA2 is the advanced and better encryption methodology for WLAN security. Moreover, combination of WPA2 and AES is the best solution for security of WLAN network. Wireless network can be made more robust by enabling MAC filtering in router setting. MAC filtering facilitates only limited users to use the network and thus an unauthorized access can be prevented. For commercial networks, network security tools can also be used for prevention of hacking as well as various cyber attacks.

Cross Layer Optimization

Communication network protocols are engineered, developed and established based on the layered approach. Each layer is designed to serve a specific functionality in collaboration with other layers. Thus, various functionalities such as transmission characteristics, error control, flow control, synchronization, routing of information, framing, sequencing of packets, congestion control, application specific services etc. are combined together by some degree of interfacing between various layers for implementation of protocols. Protocol layers are designed and organized in a vertical hierarchical manner. They are designed to ''pass on'' the particular message. Communication is possible only between adjacent layers for the purpose of responding as a part of passing on the information.

In TCP-IP and OSI models, each layer provides services to its upper layer. Layered optimization and design has been a well-established approach for communication network design and development. In layered approach, protocols are designed with independent functionalities of layers. The specific layer uses the services of lower layers irrespective of the process and parameters of the service provided. A specific layer is concerned about a layer located above or below it and that too only for the sake of limited responses and communications.

Layered approach for wired communication approach can be modified with cross layer approach for wireless communication for performance enhancement. For example, cognitive radio technology is meant to provide unused spectrum of primary users to secondary users for enhancing spectrum efficiency. This approach requires continuous monitoring of channel conditions, interference and traffic scenarios. The primitive layered approach is not suitable for cognitive radio technology in which a reliable coordination between layers and adaptation is inevitable for successful operation and management of spectrum. In the context of Smart grid technology, cognitive radio technology is an inevitable approach for gigantic communication network and enormous amount of data communication.

Moreover Smart grid technology comprises hierarchical and heterogeneous network with diverse set of communication protocols (Shah et al., 2013). This demands a divergence from primitive approach and adaptation of an innovative approach.

The cross layer approach can be defined in various ways depending upon its design and functional aspects as described as follow:

- A design approach to explore the synergy and collaboration between various network layers.
- A combined design and optimization approach by considering more than one layer.
- A collaborative approach between different layers by sharing of information between them.
- In the context of multimedia transmission, a cross layer optimization is a process of collective source and channel coding.
- A design and optimization approach by considering the connection between various layers.
- An approach to explore dependence between network layers in contrast to independent layered approach.
- A combined parameter optimization for different layers.

A cross layer approach is a desecration of traditional layered design approach through ‘’Collaborative Optimization”. For Smart grid technology, this approach is a very essence of optimization of communication network protocols. The transmission and reception characteristics are taken into consideration at higher layers and the data communication is considered at lower layer is one of the aspect of cross layer approach. Non adjacent layers are collaboratively designed and optimized for performance enhancement. Cross layer optimization can be achieved for performance enhancement through either combined optimization of parameters pertaining to various network layers or by exchange of information between various network protocol layers. Figure 3 shows the conceptual diagram of cross layer design.

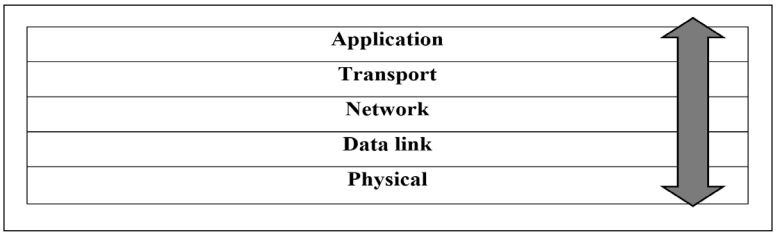


Figure 3. Conceptual diagram of cross layer design

Adhoc wireless networks play a crucial role in Smart grid communication architecture especially in home automation applications. Adhoc wireless networks work on the basis of decentralized or distributed control. This feature imposes difficulty and challenges to serve and support high data throughput and delay constrained applications which are inevitable in Smart grid. The traditional layering methodology used in wired as well as wireless

approach in which each layer is unmindful about operation of rest of the layer cannot provide optimum performance under rigorous performance and reliability requirements. The protocol layer optimization of isolated layers may have negative impact on performance and operation of other layers.

The austere performance and operation demands of Smart grid heterogeneous network can only be realized through cross layer optimization (Shakkottai et al., 2003). Cross layer design approach explores interdependencies between diverse network layers through parameter characterization and optimization to optimize an entire network performance. Cross layer design can be used for wired as well as wireless networks. For wireless networks, this approach becomes inevitable to address the challenges such as connectivity, data throughput, QoS and interference. Adaptive power control, signaling, modulation and encryption can address performance challenges. Performance optimization of physical layer parameters can be explored by higher layers to optimize the network performance. Fading and congestion at physical layer can be circumvented by higher layers with the help of adaptive routing for delay minimization. The bottlenecks resulted due to one layer can be avoided if the joint optimization is performed and the operational characteristics of one layer is shared to rest of the layers. If an information regarding data throughput and delay is known to application layer then the solutions such as changing the rate of compression and/or multiple routing can be optimized. Cross layer optimization approach imposes many design challenges as adaptation or deviations at a specific layer must be compensated at that layer in terms of time scale. The variations in the Signal to Interference Ratio (SINR) are faster in terms of microseconds. Variation in the traffic is in terms of few seconds and topology of a network takes considerable time to change. Cross layer optimization must be performed by adaptations or optimization of parameters pertaining to a specific layer. If it is ineffective then the information must be exchanged with higher layer for better response and resolution. For example, if SINR is changing at a very fast rate as a result of fading and the physical layer shares this information to upper layers for the solution of this issue then during that time, SINR will mostly change to some other value or the fading might have reduced to optimum level. So, the problem must be first addressed at that specific layer before exchanging it with other layers. Optimization at other layer must resolve or alleviate the effect of performance degradation.

If SINR is decreasing due to fading then a physical layer may resolve this problem by increasing the transmitter power. But if this problem is a result of mobility or change in the location where the signal power is weak (for example a mobile station enters into the tunnel) then it's a momentary fading which will be resolved as soon as the mobile station is out of it. This type of fading can be resolved by procrastinating the packet transmission at higher layers. For extremely itinerant nodes, the solution can be adaptations in link characteristics, change in network topology or adaptations in routing methodology. WSN

has challenging energy and delay constraints which can be solved through joint optimization at network and MAC layer. In joint or adaptive cross layer optimization method, each layer must be tuned to the variations done for performance optimization. Cross layer design is a challenging task as it is complex and requires multidisciplinary expertise. The successive part of this chapter describes various case studies of cross layer optimization.

Parameter Optimization of Neighborhood Area Network

Neighborhood Area Network (NAN) is a combination of Home Area Networks in Smart grid hierarchical network. In this section, NAN is formed in 1.5 Km area using IEEE 802.11 standard. Total 50 nodes are considered for optimization as shown in Figure 4. In this section, various parameters of different routing protocols are optimized and compared.

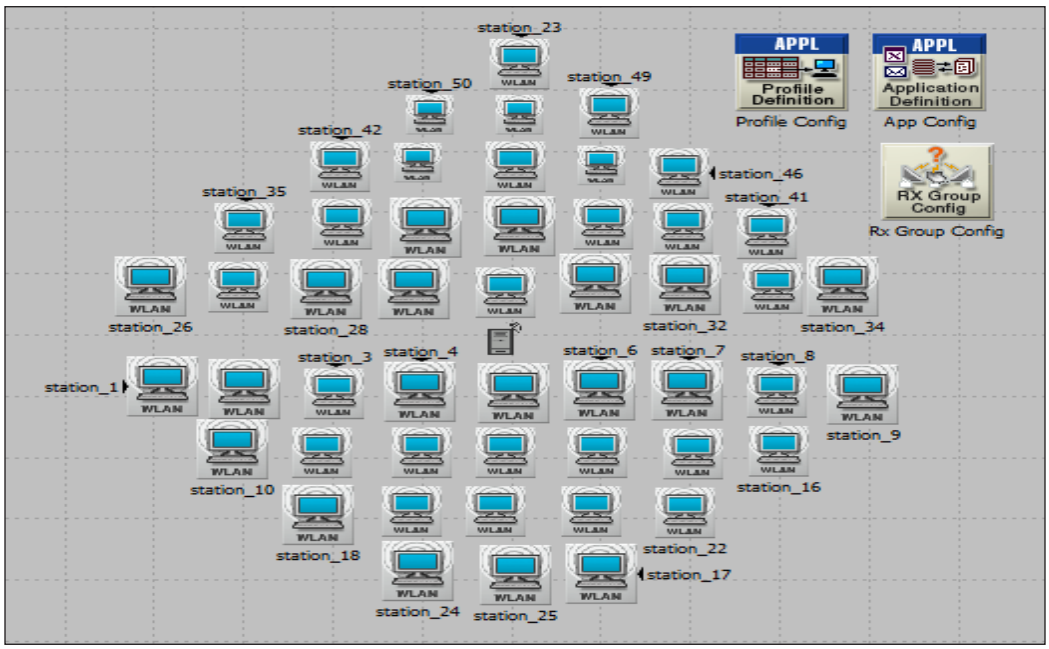


Figure 4. Neighborhood area network

Comparison of WLAN Routing Protocols

In this section, routing protocols are compared for optimization purpose. Table 2 shows the theoretical comparison of various protocols.

Default parameters are considered for all the routing protocols. Adhoc On demand Distance-Vector routing protocol works on the basis of on demand routing, that is, the route is established only when there is a requirement for source node to transmit the packets. The delay caused by set up of connections is lesser in this protocol.

Dynamic source routing protocol restricts the bandwidth consumption through elimination of table update messages. It does not use hello packet approach. Temporally Ordered Routing Algorithm protocol uses loop free multipath ways towards destination. It uses a link reversal algorithm for this purpose. It has less overhead as it restricts the control packets for reconfigurations of routes to a small region.

Optimized link state routing protocol enhances the performance of Link state routing protocol by reducing the size of control packets as well as number of links used to forward the link state packets. OLSR has many advantages such as less overhead, less connection set up time and number of broadcasts. Geographic Routing Protocol is a location based proactive protocol. It uses Global Positioning system for updating flooding information. Results show that OLSR protocol performs better in terms of total delay, media access delay and throughput compared to all other protocols.

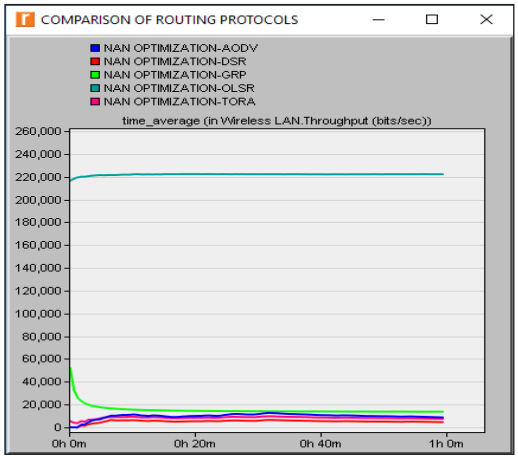


Figure 5. Comparison of throughput of different routing protocols for network optimization

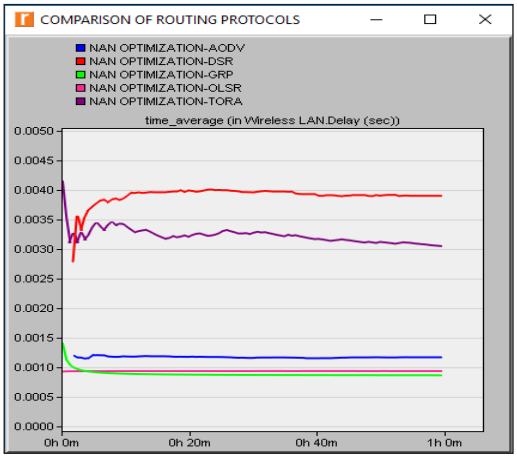


Figure 6. Comparison of WLAN delay of different routing protocols for network optimization

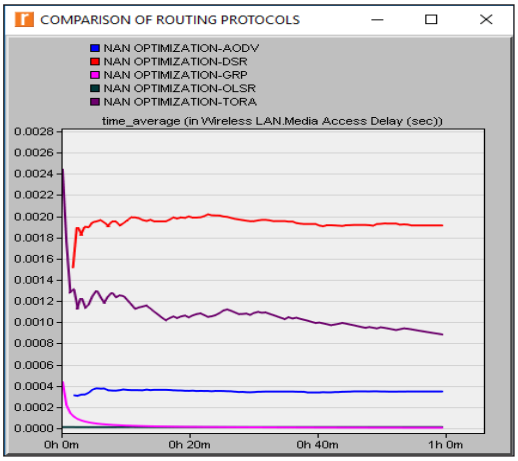


Figure 7. Comparison of WLAN Media access delay of different routing protocols for network optimization

Performance Enhancement in OLSR Protocol

As depicted in above results, OLSR protocol performs better than other routing protocols. It can be further improved by increasing Hello interval, Topology hold time and Duplicate message hold time. TC interval is decreased. The various parameters are described below. As shown in graphical results, WLAN throughput increases significantly as a result of parameter optimization. Table 2 shows the description of various parameters.

Table 2
Description of various parameters

SR NO.	Parameter	Description
1.	Willingness	This attribute defines that whether the node is agree to forward or carry the traffic on behalf of rest of the nodes or not.
2.	Hello Interval	This aspect specifies the time intermission between hello packets. Hello packets are essential to retain adjacencies between various 1 hop or 2 hop adjacent nodes.
3.	TC Interval	This feature states the interlude between TC messages. TC messages are used for routing table calculations and spread the information pertaining to topology of nodes.
4.	Neighbor hold time	This time period defines the duration within which a hello packet should be reached in order to sustain the link with neighbors. It is generally three times the hello interval.
5.	Topology hold time	This characteristic states the termination time for records in topology table. This feature is usually set to 3 times the TC Interval.
6.	Duplicate message hold time	This feature states the finishing time of an entry in the duplicate set table.
7.	Addressing mode	It specifies the protocol IPV4 or IPV6. The packets unsupported by the set parameter will be dropped.

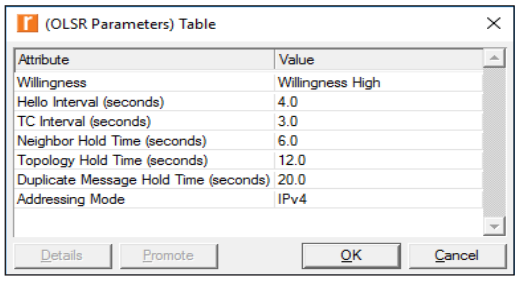


Figure 9. Optimized parameters

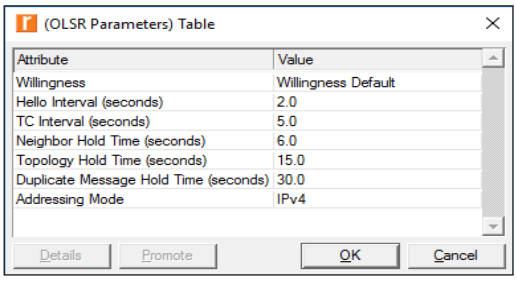


Figure 8. Default parameter

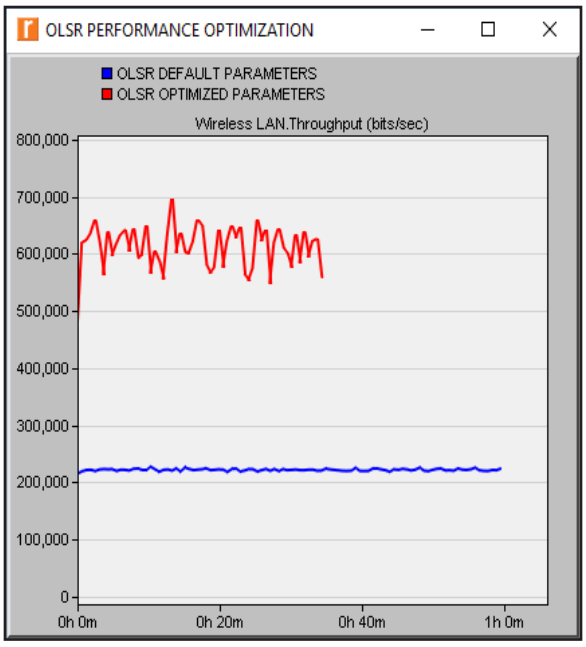


Figure 10. WLAN Throughput for default and optimized OLSR parameters

RESULTS AND DISCUSSION

In this paper, the performances of various routing protocols such as AODV, DSR, GRP, OLSR and TORA are analyzed for Neighborhood Area Network. As shown in Figure 5, the throughput is the highest for OLSR and the lowest for DSR protocol. The delay as well as media access delay is maximum for DSR and minimum for OLSR and GRP protocols as depicted in graphical results shown in Figure 6 and 7 respectively. OLSR performs better compared to rest of the protocols for implemented Adhoc network. For different applications such as high quality multimedia and VoIP, diverse set of protocols can be a suitable choice based on analysis, optimization and performance. Moreover, the network performance is further enhanced by increasing Hello interval, Topology hold time and Duplicate message hold time as shown in Figure 9. TC interval is decreased for performance enhancement. Network performance can be optimized through adaption of various parameters as shown using graphical representation in Figure 10. Thus after cross layer modifications, OLSR outperforms other protocols for Smart grid NAN applications. The paper also includes security protocols and solutions for network security. The combination of WPA2 and AES is an optimum solution for WLAN security. MAC filtering can also be implemented for prevention of unauthorized network access.

CONCLUSION

Smart grid is the most imaginative and advanced technology of present era. It is a hierarchical network comprising of heterogeneous technologies. The layered architecture of Smart grid communication network consists of Home Area Network, Neighborhood Area Network and Wide Area Network. Smart grid is a novel concept which is still in a developing stage. It is a complex and gigantic network of information and electrical infrastructures and equipment. The primitive layered approach for existing communication networks is not apt for optimized performance and operation of Smart grid network. Joint design of different protocol layers provides enhanced network performance. OLSR has average end to end delay which is suitable for applications which demands lesser delay. OLSR routes the information periodically. Sequential delivery of messages is not mandatory for OLSR which makes it independent of link reliability issue. OLSR performs better than rest of the protocols for Smart grid NAN applications but network optimization is inevitable for diverse applications such as multimedia and VoIP applications which require huge bandwidth and real time streaming of bulk data. WLAN is a ubiquities network. So, network security is the utmost priority for WLAN users. Implementation of security measures such as combinational security protocols, MAC filtering and network security tools can make the network robust against hacking attacks.

REFERENCES

Chen, B., Wu, M., Yao, S., & Ni, B. (2006). ZigBee technology and its application on Wireless Meter-reading System. In *IEEE International Conference on Industrial Informatics (INDIN 06)* (pp. 1257-1260). Singapore.

Erol-Kantarci, M., & Mouftah, H. (2011). Wireless multimedia sensor and actor networks for the next generation power grid. *Ad Hoc Networks, 9*(4), 542-551.

Farooq, H., & Jung, L. (2014). Choices available for implementing smart grid communication network. In *Computer and Information Sciences (ICCOINS)* (pp.1-5). Kuala lumpur, Malaysia.

Gungor, V., Lu, B., & Hancke, G. (2010). Opportunities and challenges of wireless sensor networks in smart grid. *IEEE Transactions on Industrial Electronics*, *57*(10), 3557-3564.

Mahmood, A., Javaid, N., & Razzaq, S. (2015). A review of wireless communications for smart grid. *Renewable and Sustainable Energy Reviews, 41*, 248-260.

Patel, A., Aparicio, J., Tas, N., Loiacono, M., & Rosca, J. (2011). Assessing communications technology options for Smart Grid applications. In *Proceedings of the 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)* (pp. 126-131). Brussels, Belgium.

Saputro, N., Akkaya, K., & Uludag, S. (2012). A survey of routing protocols for smart grid communications. *Computer Networks, 56*(11), 2742-2771.

Shah, G., Gungor, V., & Akan, O. B. (2013). A Cross-Layer QoS-Aware Communication Framework in Cognitive Radio Sensor Networks for Smart Grid Applications. *IEEE Transactions on Industrial Informatics, 9*(3), 1477-1485.

Shakkottai, S., Rappaport, T., & Karlsson, P. (2003). Cross-layer design for wireless networks. *IEEE Communications Magazine, 41*(10), 74-80.

Sun, Z., Zhao, T., & Che, N. (2009). Design of electric power monitoring system based on ZigBee and GPRS. In *1st International Symposium on Computer Network and Multimedia Technology (CNMT 2009)* (pp. 1-4). Wuhan, China.

Xu, Z., Xue, Y., & Wong, K. (2014). Recent Advancements on Smart Grids in China. *Electric Power Components and Systems, 42*(3-4), 251-261.

Pertanika J. Sci. & Technol. 26 (4): 1725 - 1749 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A New Genetic Algorithm Based Technique for Biomedical Image Enhancement

Khatkar, Kirti[1*] and Kumar, Dinesh[2]

[1]Research Scholar, Department of Computer Science & Engineering, GJUS&T, Hissar,Haryana, India
[2]Department of Computer Science & Engineering, GJUS&T, Hissar, Haryana, India

ABSTRACT

From a diagnostic perspective, image enhancement has diverse potential in image processing applications related to biomedical images. A hybrid algorithm obtained by combining discrete wavelet transformation with soft computing techniques is proposed for enhancing the biomedical images. This paper proposes an approach for effective visual enhancement of biomedical images. The proposed approach uses scale-invariant feature transform algorithm and principal component analysis as pre-enhancement steps, followed by the combination of DWT and the genetic algorithm to enhance the biomedical images. In GA, a new fitness function, which can efficiently reduce the noise in biomedical images while preserving the details, is proposed for the enhancement process. In order to accurately evaluate the enhanced image's quality, various metrics like peak signal to noise ratio, contrast to noise ratio, BETA coefficient, standard deviation, and mean square error have been considered. Finally, the comparison of the proposed algorithm with other soft computing techniques like Bacterial Foraging, Particle Swarm Optimization and Fuzzy Logic is carried out. The results show that the proposed technique outperformed over the other methods and provided better image quality.

Keywords: Bacterial foraging, denoizing, standard deviation, Fuzzy logic, genetic algorithm, Haar-wavelet, image enhancement, pre-processing

ARTICLE INFO

Article history:
Received: 20 October 2017
Accepted: 27 July 2018
Published: 24 October 2018

E-mail addresses:
kirtinitkuk@gmail.com (Khatkar, Kirti)
dinesh_chutani@yahoo.com (Kumar, Dinesh)
* Corresponding author

INTRODUCTION

Various Images like ultrasound, Computed Tomography (CT), and Magnetic Resonance Imaging (MRI) are used globally clinical diagnosis and treatment. Image quality is the key aspect while considering any biomedical image for diagnostic purposes. While capturing the biomedical images, the

ISSN: 0128-7680
e-ISSN: 2231-8526

noise introduced due to various environmental factors, distorts the captured image quality, in particular damaging the structure and the contents of the image as well as the correlation between pixels. Thus, to remove the noise various pre-processing steps like denoizing, restoration, enhancement, sharpening, and brightness correction are used (Wang, Jiang, & Ning, 2012). For improving the image visualization, enhancement is considered to be of key importance (Takeda, Farsiu, & Milanfar, 2007). However, some enhancement algorithms emphasize only on enhancement than noise reduction, like histogram equalization in wavelet domain (Fu, Lien, & Wong, 2000). Bayesian estimation for multi-component image in wavelet domain (Scheunders & De Backer, 2005), contrast enhancement using modified coarse and detail coefficients (Xiao & Ohya, 2007). Ercelebi and Koc (2006) proposed the lifting-based wavelet domain Wiener filter for noise reduction. Zeng, Dong, Chi and Xu (2004) and Jung et al. (as cited in Claudio & Scharcanski, 2004) proposed wavelet-based methods for both enhancement as well as noise suppression.

The main aim of this paper is to present an effective approach for enhancing the biomedical images for the diagnosis. The blurred biomedical images may lead to wrong diagnosis of diseases, such as lesions in breast cancer diagnosis, tumor detection in brain, and blood vessel thickening. Thus, it becomes imperative to propose a technique that can result in effective image enhancement for a better diagnosis. The advantage of the proposed approach is that it works well on all types of images like ultrasound, MRI, and CT. In the present study, all such images of different human body organs like abdomen, brain, kidney, liver, and lungs have been used. In the pre-enhancement stage, the SIFT and PCA techniques were used. In the second stage a combination of DWT with GA was used for the enhancement purpose. The following section briefly describes wavelets, which are predominantly used for noise reduction.

Wavelet and Biomedical Image Enhancement

Wavelets have been successfully incorporated in various image compression, enhancement, analysis, classification and retrieval applications. Jansene (2001) proposed a wavelet-based denoizing method by using the thresholding concept. Xie, Pierce and Ulaby (2002), and Pizurica, Philips, Lemahieu and Acheroy (2003) proposed the non-homomorphic wavelet filtering techniques for synthetic aperture radar and ultrasound images. Solbo and Eltoft (2004) used homomorphic wavelet filtering of SAR images, which essentially uses logarithmic images instead of the original images, as is the case while dealing with non-homomorphic ones. Later, Selesnick et al. (Selesnick, Baraniuk, & Kingsbury, 2005) proposed a new concept of Dual-Tree Complex Wavelet Transform (DTCWT) for image enhancement. Thavavel and Murugesan (2007) used complex wavelet transform for CT images. Bosdorf, Raupach, Flohr and Hornegger (2008) used the wavelet-based correlation

analysis for CT images to extract the uncorrelated noise component and found that the noise in CT images is non-white.

Tan and Shi (2009) and Anand and Sahambi (2010) used wavelet-based filters for the removal of rician noise that was predominant in MRI images. Later, Wang, Jiang and Xing (2012) proposed a GSM model based wavelet method for CT image denoizing. Xiang-wei, & Yu-xiu, (2015) used wavelet multi-resolution analysis method for image enhancement in which the soft thresholding was applied to obtain an enhanced image. One step further, Rasti, Daneshmand, Alisinanoglu, CagriOzcinar and Anbarjafari (2016) proposed a stationary wavelet transformation for enhancing the image. The previous techniques were used on a single type of image; whereas an effort is being made in the present paper to propose a technique, based on the combination of wavelets and soft computing, which runs well on all kinds of images.

Proposed Work

In this study, we intend to expand the previously suggested methods of enhancements. From the literature review, it is evident that most of the techniques proposed for enhancing the biomedical images are specific to either ultrasound or CT images or MRI images. The technique proposed in this paper works equally well for all kinds of images. By combining the wavelets with soft computing techniques, an effective enhancement technique is proposed for the biomedical images, which would be valuable for diagnostic purposes. This method is divided into two steps: pre-processing step and the main step. Firstly, the SIFT algorithm is implemented on the medical image. The SIFT algorithm is run for as many as hundred iterations on every image under study and the best result in terms of PSNR is considered for PCA application. PCA fully de-correlates the original dataset.

As the energy of a signal mainly lies on the PCA transformed dataset and the random noise is evenly spread over the whole dataset, the signal and noise can be better distinguished in the PCA domain. Thus, PCA is also used as a part of the pre-processing step to enhance SIFT's results.

Pre-processing step

The SIFT and the PCA methods are applied as a pre-processing step in enhancing the medical image.

SIFT (Scale Invariant Feature Transform)

SIFT can vigorously recognize objects among disordered and under incomplete occlusion (Lowe, 2004). Since the SIFT feature descriptor is invariant to uniform scaling and orientation, and partially invariant to illumination changes, it is best suited for feature

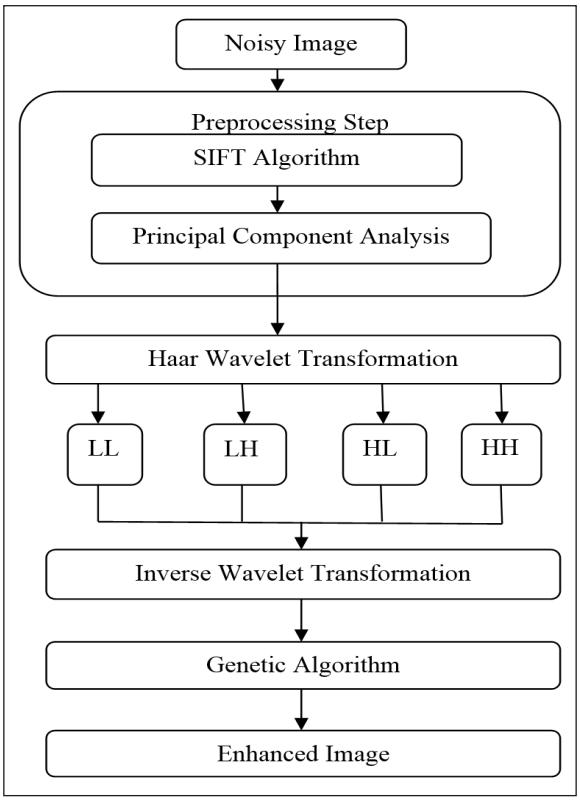


Figure1. Schematic diagram of the proposed method

matching. SIFT key points are first extracted from a set of reference images on the basis of the Euclidean distance of feature vectors. The matching features for the candidate image are extracted.

In the present work when the SIFT algorithm is applied to the noisy biomedical image; the initial image may contain key points that are of low contrast, which is removed by the algorithm. The effect of scale and rotation around each point is removed by selecting areas around every point thus providing better results. For best results, as many as hundred iterations of the SIFT algorithm have been considered individually for different biomedical images. The iteration that gives the highest value for PSNR is considered for further processing. The resultant image is then processed by PCA.

PCA (Principal Component Analysis)

Pearson (1901) proposed PCA as an illustration of the principal axis theorem which was later developed by Harold Hotelling in the 1930 (Hotelling, 1936). The basic logic of PCA is that it converts a set of correlated variables into a set of linearly de-correlated variables called principal components. By using Shannon Entropy (Geiger & Kubin, 2012), PCA reduces the amount of information lost during the dimensionality reduction.

In the proposed work, PCA is applied on the resulting image after SIFT algorithm, it divides the image dataset into principal components that contain the relevant features while the other dataset contains highly uncorrelated noise. PCA low order approximation act as a noise filter by separating correlated features from the uncorrelated noise and preserving them. Thus, it improves the PSNR value in the resulting biomedical image.

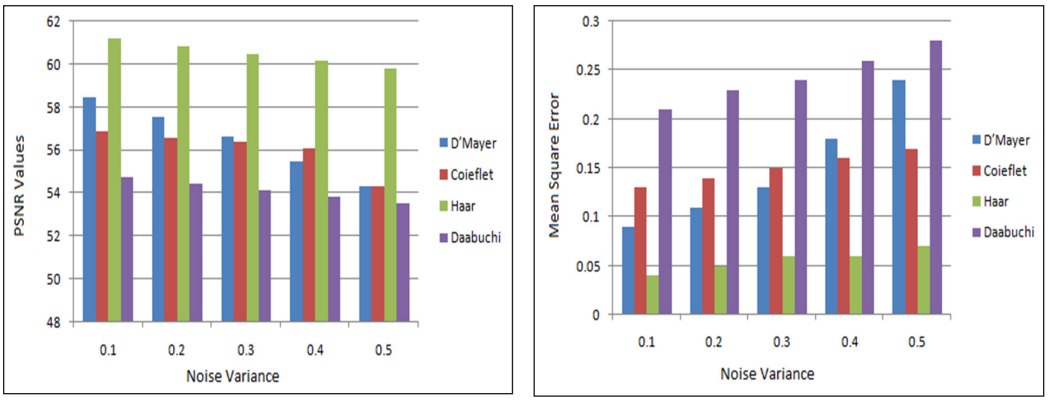
Enhancement Method

After performing the pre-processing step on the sample image the Haar wavelet is used for the image enhancement. The wavelet transformation has various properties like good image representation, multi-resolution analysis, data reparability and compaction, which make it a more powerful tool for biomedical image enhancement. In the present approach, the Haar wavelet has been used on the results obtained after applying PCA. The resulting values of various metrics considerably amount of enhance the image. Genetic Algorithm is applied to the Haar results for further optimization of the results.

Enhancement using Wavelets

The use of DWT-based techniques is a recent trend for speckle removal (Bao & Zhang, 2003; Wang, Bovik, Sheikh, & Simoncelli, 2004; Khare, Khare, Yongyeon, Hongkook, & Moongu, 2010). The Haar wavelet is the first known wavelet and was proposed by Alfred Haar in 1909. Its conceptual simplicity, memory efficiency, and reversible nature without edge effect make it more efficient than other wavelets. Due to the multiplicative nature of speckles, the Haar wavelet uses logarithmic transformation to convert multiplicative speckle noise to additive white noise. Subsequently, the wavelet thresholding is used to remove the noise. Later the reverse logarithmic transformation is done (Pizurica et al. 2003). Unlike the traditional method, a combination of wavelets with the GA is used for enhancement purpose in the proposed method. The resulting image after PCA is used as an input for the DWT.

The Haar wavelet is the shortest existing wavelet having a filter for analysis and synthesis of length two due to which the pixel coefficients that lie close to edges, are efficiently retrieved without information loss. As few pixels are used for correlation analysis, noise could be wrongly detected as structures resulting in the appearance of white spots in the Haar wavelet output image. GA is used for further image enhancement. GA has been proven to be a powerful optimization technique in a large solution space and has applications in the biomedical field. The comparison of Haar wavelet with other wavelets for PSNR and MSE is as shown in figures 2(a) and 2(b).The results show that the Haar wavelet provides a better PSNR value than other wavelets, and is hence considered in the proposed method for better results.



(a) (b)

Figure 2. Comparison of different wavelets with the Haar wavelet considering PSNR and MSE

Genetic Algorithms

Genetic algorithms (GA) were developed by John Holland on the basis of the mechanics of natural selection and natural genetics. The GAs are dependent on the concept of "Survival of Fittest". GAs uses a process consisting of selection, crossover and mutation operators. GAs follows the successive generations to choose a chromosome structure. An initial population is randomly generated. Genetic operators, such as crossover and mutation are applied to achieve the desired optimized results (Liu, 2015; Kaur, Gurvinder, & Parminder, 2016).

Steps performed during GA:

1. Initial Population: An initial population of N chromosomes, is randomly generated within the search space. In the initial population, a small number of chromosomes leads to poor results, while a large number results in greater computation time. An optimum number of chromosome population is desirable for good results. Therefore, the number of chromosomes, N, are considered as 50 in the present work and are kept constant in all generations.

2. Fitness Function: In order to evaluate the fitness of chromosomes in every generation the fitness function F_t, which is the average of the summation of each pixel value of the image matrix, is computed by using the formula:

$$F_s = \frac{1}{p}\left[\sum_{j=1}^{p} \frac{1}{q}\left[\sum_{i=1}^{q} a_{i,j}\right]\right] \quad (1)$$

where p is the number of rows and q is the number of columns in an image matrix.

3. Crossover and Mutation: Genetic crossover is a probabilistic process that exchanges information between two parent chromosomes for generating two child chromosomes.

The arithmetic crossover which produces two complementary linear combinations of the parents selected chromosomes is replaced by new chromosomes that are derived after application of genetic operators. Finally, the results are obtained by choosing the chromosome with the highest fitness value. The results of the proposed method provide better-enhanced images in terms of an increased PSNR value.

Image Quality Assessment

The distortions during acquisition, processing, compression, storage, transmission, and reproduction are responsible for degradation in biomedical images quality. According to Wang, Jiang and Xing (2012) the subjective method of image quality evaluation is inconvenient, time-consuming, and expensive. Therefore, the objective image quality assessment is mainly used in diagnostic applications using quantitative methods to predict the perceived image quality. The important features of image quality assessment like dynamic monitoring and adjusting of image quality, to optimize algorithms and the parameter settings of image processing systems, make it more appropriate for biomedical image applications (Zhang, Wang, & Duanmu, 2010). The enhanced image quality is measured by comparing: peak signal to noise ratio (PSNR), standard deviation (SD), mean square error (MSE), contrast to noise ratio (CNR) and edge detection (BETA).

PSNR of the enhanced image is compared with other images. PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the reliability of its representation. It is calculated using Equation 2. The greater the PSNR value the better the image quality.

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}_1^2}{\mathrm{MSE}}\right) = 20\log_{10}\left(\frac{\mathrm{MAX}_1}{\sqrt{\mathrm{MSE}}}\right) \tag{2}$$

The beta metric is used as edge and preservation measure in the filtered image (Beis & Lowe, 1997)

$$\beta = \frac{r(\Delta I - \bar{\Delta I}, \Delta \hat{I} - \bar{\Delta \hat{I}},)}{\sqrt{r(I - \bar{\Delta I}, \Delta I - \bar{\Delta I}).r(\Delta \hat{I} - \bar{\Delta \hat{I}}, \Delta \hat{I} - \bar{\Delta \hat{I}}),}}; (I_1, I_2) = \sum_{(i,j)\in ROI} I_1(i,j) \,.\, I_2(i,j)) \tag{3}$$

where ΔI and $\Delta \hat{I}$ represent the high pass filtered version of original image $I(i,j)$ and its denoizing version $\hat{I}(i,j)$. $\bar{\Delta I}$ and $\bar{\Delta \hat{I}}$ are the mean intensities of ΔI and $\Delta \hat{I}$ respectively. An increasing β indicates a better image quality.

Thirdly the CNR is defined as

$$\mathrm{CNR} = \frac{|\mu_d - \mu_u|}{\sqrt{0.5(\sigma_d^2 + \sigma_u^2)}} \quad (4)$$

Where μ_u and σ_u are the mean, and the SD is computed in an undesired region of interest (UROI), such as background. CNR measurements are proportional to the medical image quality.

We used the SD (the square root of variance) as an estimate of the signal contrast. An unbiased estimate in a discrete form is given by

$$\mathrm{S.D} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)^2} \quad (5)$$

MSE shows the average square error between a clean image and an image with error. The lower the MSE, the higher is the denoizing performance. It is given by:

$$MSE = \sum_{i=0}^{R}\sum_{j=0}^{C}\frac{(ls(i,j) - ld(i,j))^2}{R.C} \quad (6)$$

where, R and C are the dimensions of the image, ls is the original basic test image and ld is the denoized reconstructed image; i and j are the image size coordinates.

RESULTS

In this study, various biomedical images have been used for evaluating the performance of the proposed method. The artificial noise, Gaussian white noise, is added to the original image with a variance of 0.2. The method is compared with other methods used like Fuzzy logic, Particle Swarm Optimization (PSO), Bacterial Foraging Optimization (BFO) (Cincotti, Loi, & Pappalardo, 2001; Passino, 2002). For BFO, the number of bacteria considered in colony is 20, the number of chemotactic steps are 20, the number of reproduction steps are 20, the number of bacteria reproduction per generation is 5, and the elimination dispersion probability is considered 0.9. The below-mentioned table justifies the choice of the above values for different variables for BFO. By keeping the number of chemotactic steps as 20, the number of reproduction steps as 20, the number of bacteria reproduction per generation as 5 and, elimination dispersion probability as 0.9, and by varying the value of the number of bacteria, the PSNR values are as depicted in Table 1.

The PSNR value is the maximum for the bacterial population of 20 and, is therefore considered. Similarly, by keeping the number of bacteria as 20, the number of reproduction steps as 20, the number of bacteria reproduction per generation 5, and the elimination dispersion probability as 0.9 and by varying the value of the number of chemotactic steps, the PSNR values are as depicted in Table 2.

Table 1
Comparison of PSNR values for different population of bacteria for Liver Ultrasound Image

No. of bacteria	10	20	30	40	50
PSNR	65.78103	65.98231	65.28624	65.45401	65.31325

Table 2
Comparison of PSNR values for different chemotactic step for Liver Ultrasound Image

No. of chemotactic step	10	20	30	40	50
PSNR	65.90456	65.98231	65.02596	65.93501	65.70984

The PSNR value is the maximum for the number of chemotactic steps 20, and is therefore considered. Similarly, for different values of elimination-dispersal probability, the reproduction steps and the number of bacteria per generation are depicted in Tables 3, 4 and 5.

Table 3
Comparison of PSNR values for different elimination-dispersal probability for Liver Ultrasound Image

Elimination-dispersal probability	0.9	0.8	0.7	0.6	0.5
PSNR	65.98231	65.35939	65.6964	65.122076	65.30377

Table 4
Comparison of PSNR values for different reproduction step for Liver Ultrasound Image

No. reproduction step	10	20	30	40	50
PSNR	65.92724	65.98231	65.75573	65.80399	65.91268

Table 5
Comparison of PSNR values for different no. of bacteria per generation for Liver Ultrasound Image

No. of bacteria per generation	5	10	15	20	25
PSNR	65.98231	65.1936	65.23485	65.87001	65.45124

From the above tables, it is evident that the best result for BFO is obtained by considering the number of bacteria in the colony as 20, the number of chemotactic steps as 20, the number of reproduction steps as 20, the number of bacteria reproduction per generation as 5, and the elimination dispersion probability as 0.9. Similarly for PSO, the population is initially considered as 100, the number of iterations is 50, and the inertia is 1.0. The results of previous studies show that the wavelets have proven to be an efficient tool for the biomedical image enhancement (Healy, &Weaver, 1991; Sattar, Floreby, Salomonsson, & Lovstrom, 1997).

In the proposed method the various image modalities like MRI, CT, and ultrasound are used for different body organs. MRI imaging uses radio waves and magnets to form an image from inside the body. MRI imaging does not use radiations and produces a greater value for the soft tissues contrast, and is therefore recommended by doctors for the diagnosis of injuries in the brain and heart. The CT scan images, on the other hand, use X-rays to produce cross-sectional images of organs in case of internal injuries. Especially in emergency rooms, as it takes less time to capture CT images. Ultrasound is a radiation free and cheaper technique when compared to CT and MRI imaging that is mainly used for body organs containing less bony structures. In the present study, images of various body organs like heart, brain, abdomen, ovary, pancreas, liver, hepatic and lungs are taken from the University of California, Irvine (UCI) repository dataset and the Science Direct data.

The proposed method with the combination of GA technique offers a considerably improved enhancement capability as compared to the conventional enhancement methods, such as the fast fourier transformation method, the conventional wavelet-based method, and the conventional exponential-Type wavelet coefficient mapping method. The results of various steps of the proposed method for different biomedical images are in Figure 3.

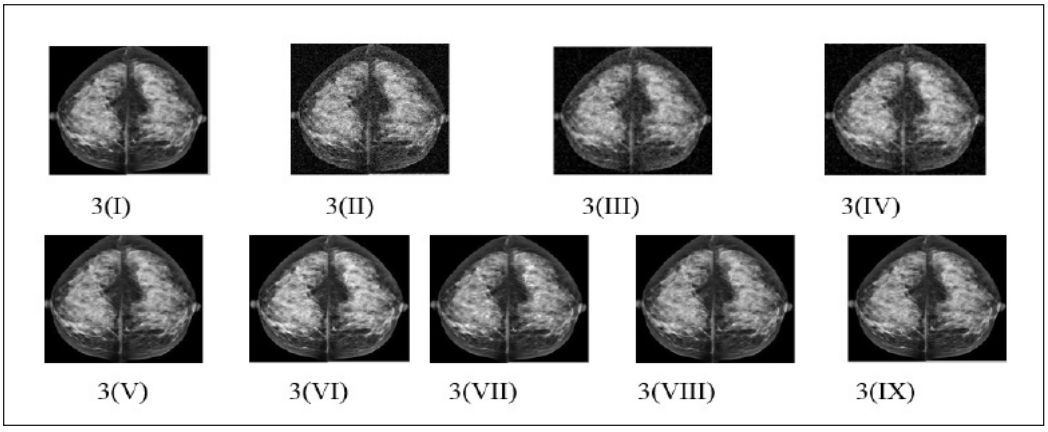


Figure 3 .Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for Breast MRI image

The other image considered for study is CT abdomen image shown in Figure 4.

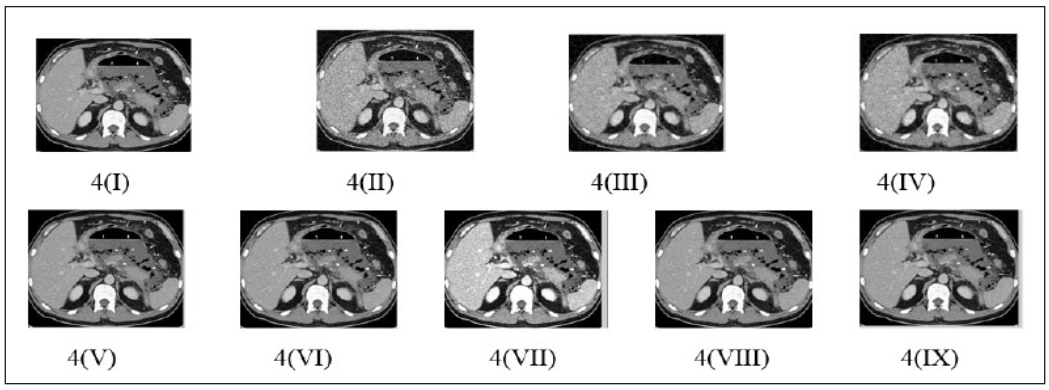


Figure 4. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for CT Abdomen image

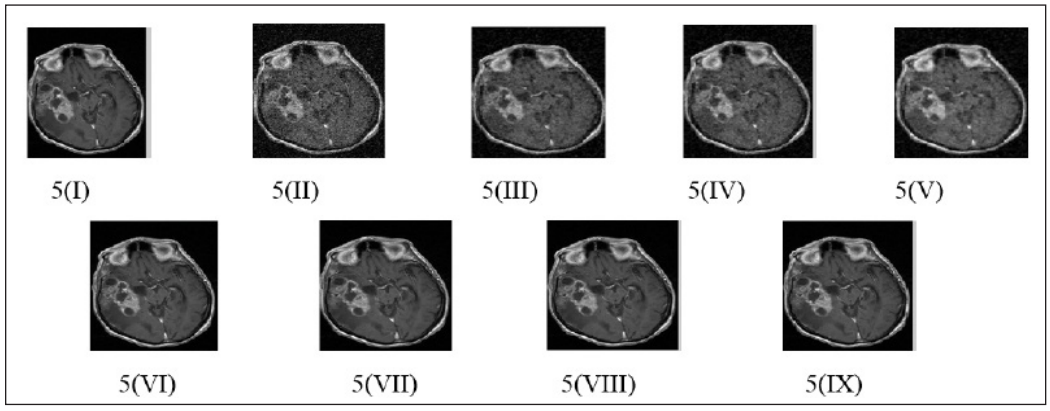


Figure 5. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for MRI Brain image

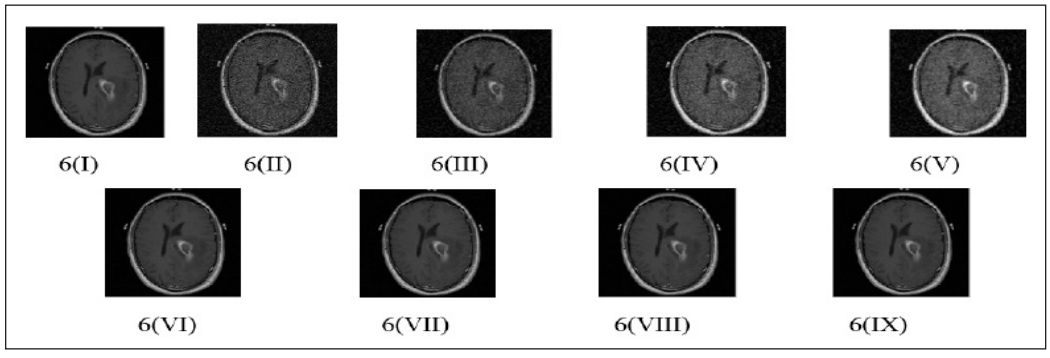


Figure 6. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for MRI Brain image

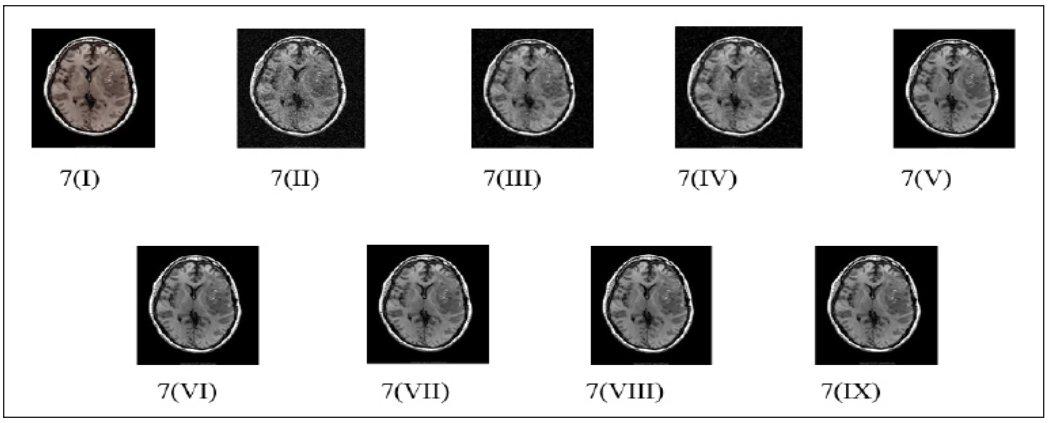


Figure 7. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for CT Brain image

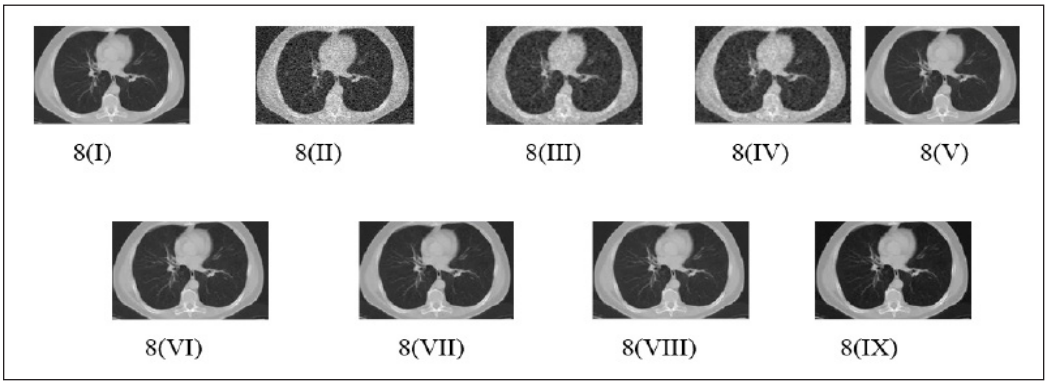


Figure 8. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for CT Lungs image

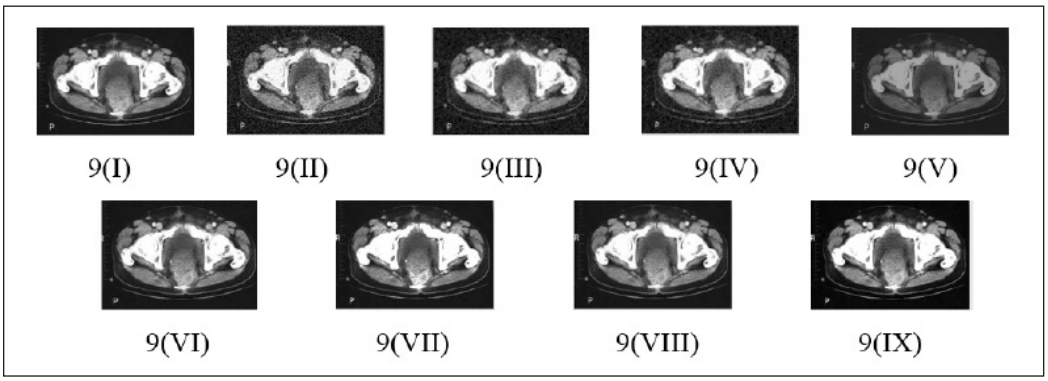


Figure 9. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for CT Rectal image

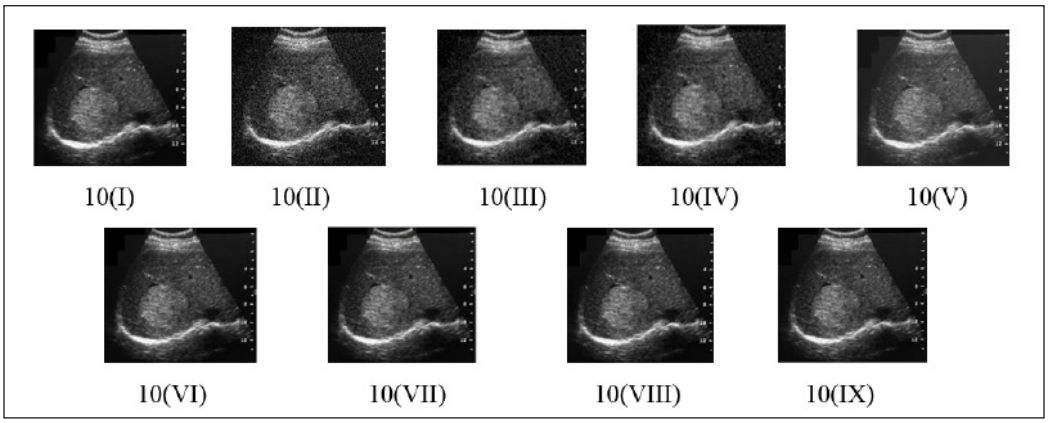


Figure 10. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for Hepatic MRI image

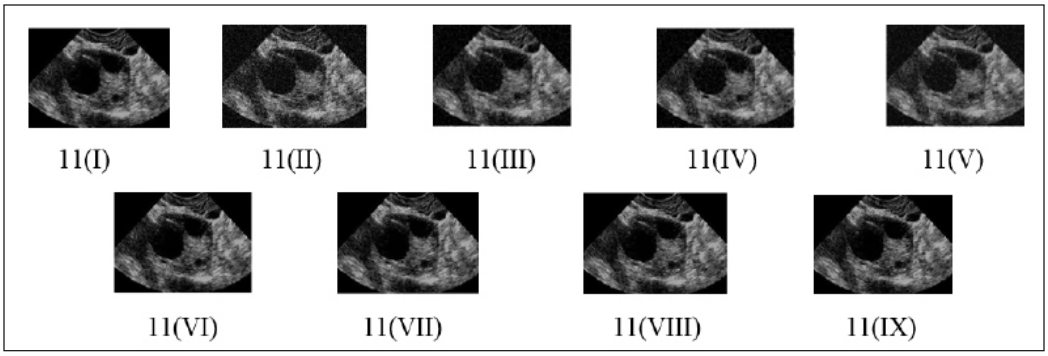


Figure 11. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for Ovary Ultrasound image

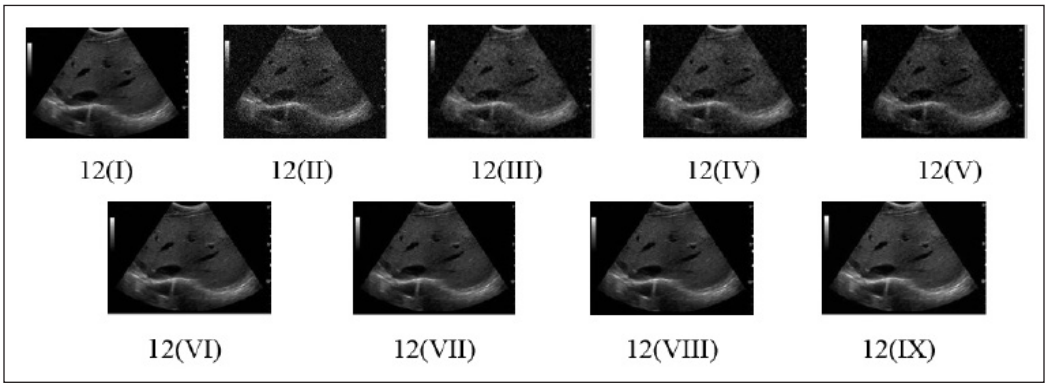


Figure 12. Results of proposed method (I) original image (II) noisy image (III) SIFT image *(IV) PCA* image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for Pancreas Ultrasound image

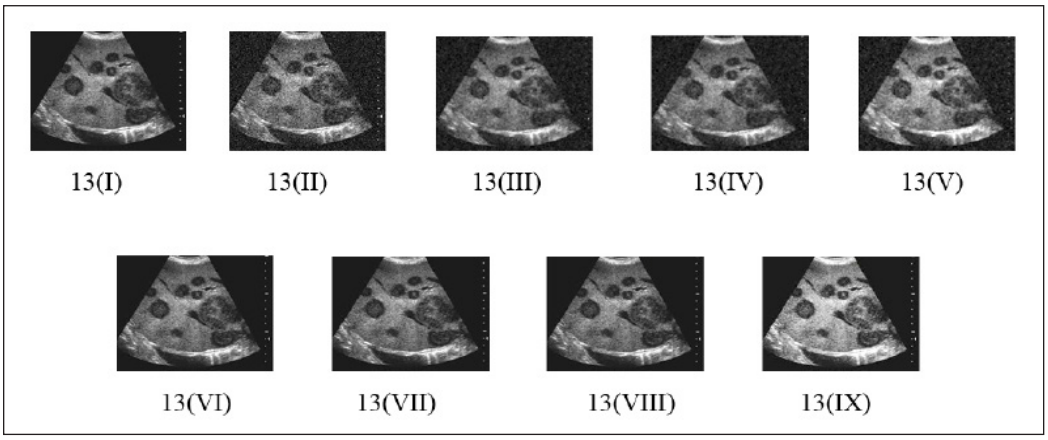


Figure 13. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for Liver Ultrasound image

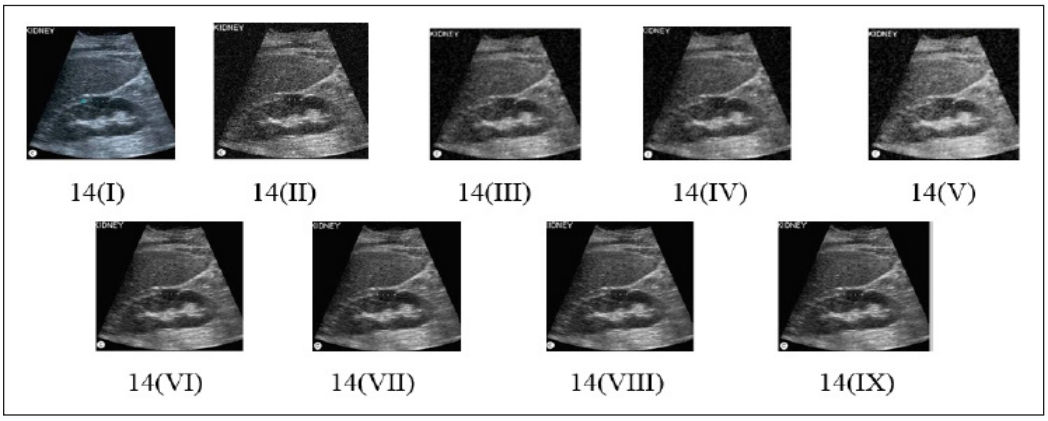


Figure 14. Results of proposed method (I) original image (II) noisy image (III) SIFT image (IV) PCA image (V) HAAR image (VI) Fuzzy Image (VII) PSO Image (VIII) BFO Image (IX) Resulting image using proposed method for Kidney Ultrasound image

In the present study, the hybrid approach is applied on different biomedical images. Figures 3 to 14 illustrate the results of the proposed method and of the Fuzzy, BFO and PSO techniques, at different stages. Tables 6 to 17 represent the comparison of different steps followed in the proposed method and illustrates the significance of each technique used in the pre-processing step, and in the main method. On the basis of the obtained results, it is concluded that the combination of wavelet and GA along with SIFT and PCA can be used as a pre-processing step, and that they can be used as an effective method for biomedical image enhancement. Hence in the proposed method a combination of these techniques has been used for the best results.

Table 6

Comparison of different steps of proposed method for image in figure 3(Breast MRI)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	61.6255	0.4472	1.5315	0.7047	2.4532
Noisy image +SIFT +GA	63.1289	0.0316	1.3037	1.9436	1.6792
Noisy image +PCA +GA	60.1186	0.6632	0.8149	1.9324	1.8791
Noisy image +HAAR+ GA	71.5508	0.0045	0.3350	1.0761	2.1716
Noisy image + SIFT +PCA +HAAR+GA	74.1779	0.0024	1.5692	1.7063	1.9132

Table 7

Comparison of different steps of proposed method for image in figure 4(CT Abdomen)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	58.2867	0.0964	1.1000	1.2934	1.5307
Noisy image +SIFT +GA	61.1764	0.0495	1.0571	1.9400	1.9722
Noisy image +PCA +GA	58.1661	0.0991	0.0857	1.0135	2.2490
Noisy image +HAAR+ GA	68.1661	0.0099	0.8745	0.7091	2.3735
Noisy image+ SIFT+PCA +HAAR+GA	70.2943	0.0060	0.5575	1.6231	2.1712

Table 8

Comparison of different steps of proposed method for image in figure 5(MRI Brain)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	60.4985	0.05647	0.7813	0.1352	1.8142
Noisy image +SIFT +GA	64.1854	0.02574	1.7875	0.2912	2.3179
Noisy image +PCA +GA	63.5120	0.02828	1.6705	0.4782	1.6356
Noisy image +HAAR+ GA	69.8745	0.00931	2.1132	0.1973	1.7452
Noisy image +SIFT+PCA +HAAR+GA	74.5706	0.0022	1.1524	1.6523	1.8867

Table 9

Comparison of different steps of proposed method for image in figure 6(Brain MRI)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	62.9897	0.04817	0.8579	0.1255	1.9192
Noisy image +SIFT +GA	65.8142	0.01978	1.8256	0.3197	2.4243
Noisy image +PCA +GA	64.1595	0.02105	1.7254	0.5288	1.7898
Noisy image +HAAR+ GA	70.8232	0.00891	2.2531	0.2173	1.8168
Noisy image+ SIFT+ PCA+ HAAR+GA	74.78	0.0021	1.0000	0.9123	1.2833

Table 10

Comparison of different steps of proposed method for image in figure 7(CT Brain)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	61.0483	0.0797	0.83954	1.9869	2.2719
Noisy image +SIFT +GA	64.1727	0.0648	1.3126	0.77882	2.0186
Noisy image +PCA +GA	61.1624	0.0787	1.7638	0.8535	1.6701
Noisy image +HAAR+ GA	68.5885	0.0151	1.6219	1.5406	1.9458
Noisy image+ SIFT+PCA +HAAR+GA	71.0483	0.0051	1.3006	0.9556	1.3436

Table 11

Comparison of different steps of proposed method for image in figure 8(CT Lungs)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	58.0538	0.1004	1.6308	2.4277	2.3097
Noisy image +SIFT +GA	61.1210	0.0923	1.754	2.0268	2.471
Noisy image +PCA +GA	58.1107	0.1017	1.8585	0.6978	2.2396
Noisy image +HAAR+ GA	68.0538	0.0101	1.6488	0.6978	2.565
Noisy image+ SIFT+ PCA +HAAR+GA	70.2041	0.0062	0.66034	1.3646	1.4227

Table 12

Comparison of different steps of proposed method for image in figure 9 (CT Rectal)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	58.9750	0.0823	0.9612	1.7394	1.9848
Noisy image +SIFT +GA	62.0558	0.0405	1.0254	1.9084	2.0359
Noisy image +PCA +GA	59.0455	0.0810	1.6129	2.0512	2.7698
Noisy image +HAAR+ GA	68.9750	0.00823	1.8612	0.9905	1.5355
Noisy image +SIFT +PCA +HAAR+GA	71.8384	0.0042	1.9019	1.9046	1.9384

Table 13

Comparison of different steps of proposed method for image in figure 10(Hepatic MRI)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	61.2160	0.0991	1.5548	2.2252	1.6809
Noisy image +SIFT +GA	64.3450	0.0639	1.0024	0.8737	2.029
Noisy image +PCA +GA	61.3377	0.0878	1.6407	1.7038	2.1821
Noisy image +HAAR+ GA	68.4021	0.0093	1.3917	0.7334	1.9339
Noisy image+ SIFT+PCA +HAAR+GA	71.2160	0.0049	1.1716	0.8736	2.289

Table 14

Comparison of different steps of proposed method for image in figure 11 (Ovary Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	58.3367	0.0953	0.4574	1.4241	2.0849
Noisy image +SIFT +GA	62.0152	0.0408	0.8794	0.5477	2.3699
Noisy image +PCA +GA	59.0049	0.08176	1.1855	2.1599	2.0183
Noisy image +HAAR+ GA	68.9391	0.0083	0.7410	0.9701	1.7159
Noisy image+ SIFT+PCA +HAAR+GA	71.7615	0.0043	0.4880	1.4585	2.4271

Table 15

Comparison of different steps of proposed method for image in figure 12(Pancreas Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	63.9420	0.0262	1.0638	2.1863	2.1191
Noisy image +SIFT +GA	66.6968	0.0139	0.9293	2.2029	1.2830
Noisy image +PCA +GA	63.6863	0.0278	0.6203	1.6225	2.5112
Noisy image +HAAR+ GA	70.8775	0.0029	0.8984	1.6975	2.4616
Noisy image+ SIFT+PCA+ HAAR+GA	73.4844	0.0010	1.7865	2.3842	2.1429

Table 16

Comparison of different steps of proposed method for image in figure 13 (Liver Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	59.8269	0.9676	0.4313	1.8273	2.0343
Noisy image +SIFT +GA	62.9231	0.5033	1.3617	0.8567	2.7824
Noisy image +PCA +GA	59.9128	0.9998	0.0231	0.5617	2.8226
Noisy image +HAAR+ GA	69.8269	0.0067	1.5643	1.4249	2.1231
Noisy image +SIFT +PCA +HAAR+GA	73.6720	0.0027	1.1984	1.9109	2.6094

Table 17

Comparison of different steps of proposed method for image in figure 14 (Kidney Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
Noisy image + GA	60.9252	0.8672	1.4699	2.1837	1.6210
Noisy image +SIFT +GA	63.5858	0.4135	0.0896	1.5606	1.7721
Noisy image +PCA +GA	60.5755	0.8998	0.9130	0.4684	2.1451
Noisy image +HAAR+ GA	68.4757	0.0058	0.5896	1.4498	2.5915
Noisy image+ SIFT+PCA+ HAAR+GA	70.4757	0.0019	0.3986	1.0382	2.3327

Tables 6 to 17 represent the numerical results of various evaluation metrics like PSNR, MSE, CNR, S.D and Beta. The above tables justify the combination of different techniques; namely SIFT, PCA, Haar Wavelet and GA in the proposed method. The result of the proposed method has been compared with other soft computing techniques likes, Fuzzy Logic, BFO and PSO. On the basis of the observed values it is concluded that the proposed method provides a better enhancement for biomedical images than other methods. The proposed method provides the best PSNR values for different biomedical images considered under the study. For the first image in table 18, the PSNR value with the proposed method is 74.1779, which is higher than the BFO, Fuzzy and PSO methods, and the method provides the least mean square value of 0.024.The value of CNR, SD and Beta coefficients are also considerably better than Fuzzy Logic and PSO, but slightly less than BFO. In other words the proposed method provides better results than the Fuzzy Logic, PSO, and BFO methods. Tables 18 to 29 also proved that the proposed method provided better results than Fuzzy Logic, PSO and BFO methods.

Table 18

Comparison of different methods for image in figure 3 (Breast MRI)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	65.6237	0.0724	1.9185	1.978	2.3496
PSO	65.7429	0.0368	0.3515	1.716	1.5058
FUZZY LOGIC	66.3927	0.0720	1.5324	0.4556	1.4768
PROPOSED METHOD	74.1779	0.0024	1.6692	1.7063	1.9132

Table 19

Comparison of different methods for image in figure 4 (CT abdomen)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	66.9059	0.1126	1.1521	1.5798	2.7222
PSO	66.6818	0.0932	0.3677	1.1188	2.1685
FUZZY LOGIC	65.8964	0.01753	0.8308	0.4907	2.4058
PROPOSED METHOD	70.2943	0.0060	0.9575	1.3231	2.6712

Table 20

Comparison of different methods for image in figure 5 (Brain MRI)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	66.9999	0.4904	1.7024	0.8362	2.7023
PSO	66.1652	0.0459	0.8373	0.2409	1.6647
FUZZY LOGIC	65.6200	0.0622	0.2080	1.4625	1.5609
PROPOSED METHOD	74.7818	0.0021	1.8477	0.8075	1.9913

Table 21

Comparison of different methods for image in figure 6 (Brain MRI)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	66.8084	0.0588	1.8643	2.3043	2.2736
PSO	66.9261	0.0584	0.6728	1.3759	1.5933
FUZZY LOGIC	66.9046	0.0586	0.1147	2.4150	2.1102
PROPOSED METHOD	74.7780	0.0021	1.8176	1.2337	1.9753

Table 22

Comparison of different methods for image in figure7 (CT Brain)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	65.8392	0.0572	1.3553	0.9526	2.8117
PSO	65.08697	0.0572	0.8639	0.1822	2.0145
FUZZY LOGIC	65.3856	0.0618	0.9267	0.3819	1.4812
PROPOSED METHOD	71.04830	0.0051	1.3006	0.9456	1.3436

Table 23

Comparison of different methods for image in figure 8 (CT Lungs)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	65.2845	0.0923	1.4720	1.2948	2.0162
PSO	65.9062	0.0815	0.4619	0.9737	2.6571
FUZZY LOGIC	66.7110	0.0134	0.9149	0.9964	1.5289
PROPOSED METHOD	70.2041	0.0062	0.9984	1.1646	1.4227

Table 24

Comparison of different methods for image in figure 9 (CT Rectal)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	66.6710	0.1628	1.4452	1.5004	1.4582
PSO	65.0926	0.0763	0.5450	0.6918	2.0368
FUZZY LOGIC	65.4779	0.0468	0.5568	0.7417	1.7675
PROPOSED METHOD	71.8384	0.0042	1.0019	0.9046	1.9384

Table 25

Comparison of different methods for image in figure 10 (Hepatic MRI)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	66.0539	0.0463	0.9778	0.9269	2.7486
PSO	66.3980	0.0051	0.5579	0.0195	2.2353
FUZZY LOGIC	65.2286	0.0113	0.6796	0.3067	1.2499
PROPOSED METHOD	71.2160	0.0049	0.9716	0.8736	2.289

Table 26

Comparison of different methods for image in figure11 (Ovary Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	66.0510	0.0860	1.8832	1.8140	2.6102
PSO	66.0328	0.0931	0.3853	0.1225	2.3759
FUZZY LOGIC	65.4883	0.1034	0.3939	0.7295	2.7056
PROPOSED METHOD	71.7615	0.0043	0.4880	1.4585	2.4271

Table 27

Comparison of different methods for image in figure 12 (Pancreas Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	65.9646	0.2864	1.7865	1.5816	1.9124
PSO	66.4595	0.0327	0.3139	0.8875	2.9031
FUZZY LOGIC	65.5932	0.3689	0.4648	0.9263	2.1423
PROPOSED METHOD	73.4844	0.0010	1.4865	1.3842	2.1429

Table 28

Comparison of different methods for image in figure 13 (Liver Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	65.9823	0.2822	1.4707	1.9006	2.1063
PSO	66.1653	0.0748	0.5245	0.1406	1.8484
FUZZY LOGIC	65.6278	0.3965	0.7342	0.7807	2.4799
PROPOSED METHOD	73.6720	0.0027	1.1984	1.1109	2.6094

Table 29

Comparison of different methods for image in figure 14 (Kidney Ultrasound)

Method	PSNR	MSE	CNR	S.D	BETA
BFO	65.1001	0.2861	1.1367	1.2161	2.6378
PSO	66.0974	0.0653	0.3457	0.5464	2.1403
FUZZY LOGIC	65.0460	0.3186	0.4432	0.6858	2.0307
PROPOSED METHOD	70.4757	0.0019	0.9986	1.0382	2.3327

Figures 3 to 14 show the results obtained by using the proposed method and by using the Fuzzy Logic, PSO and BFO methods. The images processed by the proposed method show the highest ranking. Tables summarize the quantitative evaluation results for the proposed method and other published methods in terms of PSNR, CNR, S.D, MSE and beta metrics. As described in the above section the PSNR, CNR, S.D and beta measurements are proportional to the medical image quality. It is evident from the tables that the images processed by the proposed method give the best results. The pixel curve for the original image and the pixel curve of the image after using the proposed method are shown below for biomedical images. The x-axis represents the number of pixels and the y-axis represents the pixel values of the images. It is evident from Figure 15 that the pixel-value profile of the image processed by the proposed method is more enhanced at the edges than that of the original image. It is also apparent from the figure that the noise has been significantly reduced by employing the proposed method.

DISCUSSION AND CONCLUSION

In this study, we proposed an algorithm which combined the wavelet and GA. The results of the evaluation, as illustrated in Figure 15.

Figures 3 to 14, suggest that the proposed method is significantly superior to the other methods. It is apparent that the proposed method combines the advantages of the two methods: denoizing and contrast enhancement. The results of the quantitative evaluation also reveal that the proposed method outperformed over the other methods. The main advantage of the proposed approach is that it works on all types of images like CT, MRI and Ultrasound images etc. Although the proposed method provides better PSNR and MSE but the value of beta coefficient is not satisfactory. Secondly, the proposed method is a time consuming process and is very complex. Hence, the proposed method can be simplified in such a way that the complexity and execution time can be reduced in future.

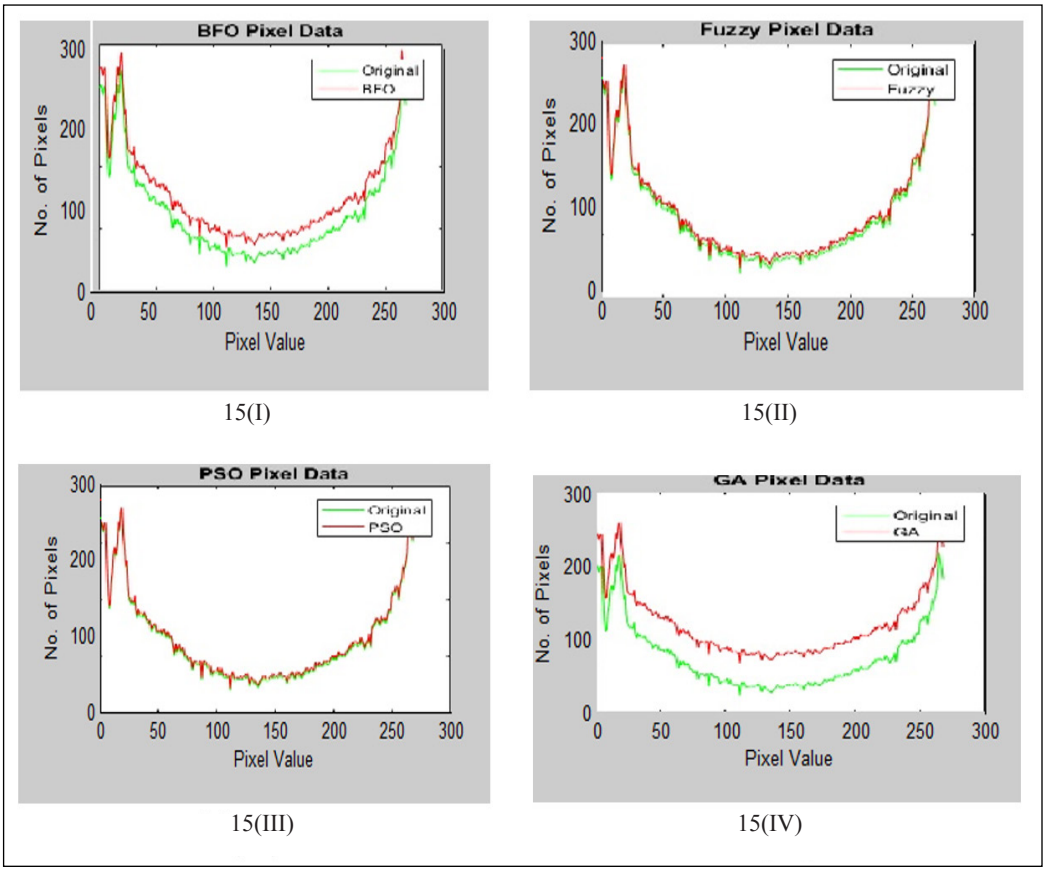


Figure 15. Pixel curve for the biomedical images using original image and Fuzzy Logic, BFO, PSO and with Proposed Method respectively for CT Rectal Image.

REFERENCES

Anand, C. S., & Sahambi, S. (2010). Wavelet domain non-linear filtering for MRI Denoising. *Magnetic Resonance Imaging, 28*(6), 842–861.

Bao, P., & Zhang, L. (2003). Noise reduction for magnetic resonance images via adaptive multiscale products thresholding. *IEEE Transactions on Medical Imaging, 22*(9), 1089–1099.

Beis, J., & Lowe, D. (1997). Shape indexing using approximate nearest -neighbor search in high dimensional spaces. In *Proceedings of the Interational Conference on Computer Vision and Pattern Recognition (CVPR97)* (pp. 1000–1006). Puerto Rico.

Bosdorf, A., Raupach, R., Flohr, T., & Hornegger, J. (2008). Wavelet Based Noise Reduction in CT-Images Using Correlation Analysis. *IEEE Transactions on Medical Imaging, 27*(12), 1685-1703.

Cincotti, G., Loi, G., & Pappalardo, M. (2001). Frequency decomposition and compounding of ultrasound medical images with wavelet packets. *IEEE Transactions on Medical Imaging, 20*(8),764–771.

Claudio, R., & Scharcanski, J. (2004). Wavelet transforms approach to adaptive image denoising and enhancement. *Journal of Electronic Imaging, 13*(2), 278–285.

Ercelebi, E., & Koc, S. (2006). Lifting-based wavelet domain adaptive Wiener filter for image enhancement. *IEEE Proceedings on Vision, Image and Signal Processing, 153*(1), 31–36.

Fu, J. C., Lien, H., & Wong, S. T. C. (2000). Wavelet-based histogram equalization enhancement of gastric sonogram images. *Computerized Medical Imaging and Graphics, 24*(2), 59-68.

Geiger, B., & Kubin, G. (2012). Relative information loss in the PCA. In *Information Theory Workshop (ITW)* (pp. 562-566). Lausanne, Switzerland.

Healy, D. M. Jr., & Weaver, J. B. (1995). Two applications of wavelets and related techniques in medical imaging. *Annals of Biomedical Engineering, 23*(5), 637–65.

Hotelling, H. (1936). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417-441.

Jansene, M. (2001). *Noise reduction using wavelets thresholding.. Lecturer series in statistics.* New York: Springer- Verlag.

Kaur, P., Kaur, G., & Parminder (2016). Image enhancement of ultrasound Images using multifarious denoising Filters and GA. In *International Conference on Advances in Computing, Communications and Informatics* (pp. 2375-2384). Jaipur, India.

Khare, A., Khare, M., Yongyeon, J., Hongkook, K., & Moongu, J. (2010). Despeckling of medical ultrasound images using daubechies complex wavelet transform. *Journal of Signal Processing, 90*(2), 428–439.

Lowe, D. (2004). Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image. *U.S. Patent No. 6,711,293.* Washington, DC: U.S. Patent and Trademark Office.

Liu, Y. (2015). Image denoising method based on threshold, wavelet transformation and genetic algorithm. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 8*(2), 29-40.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine, 2*(11), 559–572.

Passino, K. M. (2002). Biomimicry of bacterial foraging for distributed optimization and control. *IEEE Control Systems, 22*(3), 52-67.

Pizurica, A., Philips, W., Lemahieu, I., & Acheroy, M. (2003). A versatile wavelet domain noise filtration technique for medical imaging. *IEEE Transactions on Medical Imaging, 22*(3), 323–331.

Rasti, P., Daneshmand, M., Ozcinar, C., & Anbarjafari, G. (2016). Medical image illumination enhancement and sharpening by using stationary wavelet transform. In *24th international conference on Signal Processing and Communication Application Conference (SIU)* (pp. 153-157). Zonguldak, Turkey.

Sattar, F., Floreby, L., Salomonsson, G., & Lovstrom, B. (1997). Image enhancement based on a nonlinear multi- scale method. *IEEE Transactions on Image Processing, 6*(6), 888–895.

Scheunders P., & De Backer, S. (2005). Wavelet-based enhancement of remote sensing and biomedical image series using an auxiliary image. In *Proceedings of SPIE—the international Society for Optical Engineering* (pp. 600105-07). Boston, MA, United States.

Selesnick, W., Baraniuk, R. G., & Kingsbury, N. G. (2005). The Dual-Tree Complex Wavelet Transform. *IEEE Signal Processing Magazine, 22*(6), 123-151.

Solbo, S., & Eltoft, T. (2004). Homomorphic wavelet –based statistical despeckling of SAR images. *IEEE Transactions on Geoscience, Remote Sensing, 42*(4), 711-721.

Takeda, H., Farsiu, S., & Milanfar, P. (2007). Kernel regression for Image Processing and Reconstruction. *IEEE Transactions on Image Processing, 16*(2), 349–366.

Tan, L., & Shi, L. (2009). Multiwavelet-Based Estimation for Improving Magnetic Resonance Images. In *2nd International Congress on Image and Signal Processing* (pp. 1-5). Tianjin, China.

Thavavel, V., & Murugesan, R. (2007). Regularized Computed Tomography using Complex Wavelets. *International Journal of Magnetic Resonance Imaging, 1*(1), 27-32.

Wang, L., Jiang, N., & Xing, N. (2012). Research on medical image enhancement algorithm based on GSM model for Wavelet coefficients. *Physics Procedia, 33*, 1298-1303.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from Error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*(4), 600–612.

Xiao, D., & Ohya, J. (2007). Contrast enhancement of color images based on wavelet transform and human visual system. In *Proceedings of the International Conference on Graphics and Visualization in Engineering* (pp. 58–63). Florida, USA.

Xie, H., Pierce, L., & Ulaby, F. T. (2002). Despeckling SAR images using a low –complexity wavelet denoising process**.** In *Geoscience and Remote Sensing Symposium* (pp. 321-324). Toronto, Ontario, Canada.

Xiang-wei, L., & Yu-xiu, K. (2015). A novel medical image enhancement method based on wavelet multi-resolution analysis. In *8th International Conference on Biomedical Engineering and Informatics* (pp. 727-731). Shenyang, China.

Zeng, P. X., Dong, H. Y., Chi, J. N., & Xu, X. H. (2004). An approach for wavelet based image enhancement. In *IEEE International Conference on Robotics and Biomeics* (pp. 574–577). Shenyang, China.

Zhang, C., Wang, X., & Duanmu, C. (2010). Adaptive typhoon cloud image enhancement using genetic algorithm and non-linear gain operation in undecimated wavelet domain. *Journal of Engineering Applications of Artificial Intelligence, 23*(1), 61–73.

Pertanika J. Sci. & Technol. 26 (4): 1751 - 1763 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Simplified a Posteriori Probability Calculation for Binary LDPC Codes

Mostari, L.[1]* and Taleb-Ahmed, A.[2]

[1]Department of Electronics, Faculty of Technology, Hassiba Benbouali University of Chlef, Algeria
[2]LEMN DOAE UMR CNRS 8520, University of Valenciennes and Hainaut Combrésis, Valenciennes, France

ABSTRACT

The LDPC (Low-Density Parity-Check) decoder must operate soft decisions calculated using: LLR (Log-Likelihood Ratio) or APP (A Posteriori Probability) according to the decoding algorithm used. The exact calculation of these decisions for high order constellations involves complicated operations. In this work, a method to simplify the APP calculation is introduced. It is programmed to adapt as perfectly as possible the transmission system to the channel type in question. This method leads to simplify the implementation of the transmission system. Simulation results show that, under the Gaussian channel, the simplified APP algorithm for 16-QAM achieves the same performance that obtained with the exact APP, while for 64-QAM we have a small performance degradation. The same simplified APP algorithm that used for the Gaussian channel can be applied, with minor operations added, for Rayleigh channel, and it shows a small performance loss with respect to the exact APP.

Keywords: A Priori Probability (APP), Binary LDPC code, Gaussian channel, Log-Likelihood Ratio (LLR), Rayleigh channel, Square-QAM-Gray mapping

ARTICLE INFO

Article history:
Received: 26 October 2017
Accepted: 19 June 2018
Published: 24 October 2018

E-mail addresses:
l.mostari@univ-chlef.dz (Mostari, L.)
Abdelmalik.Taleb-Ahmed@univ-valenciennes.fr (Taleb-Ahmed, A.)
* Corresponding author

INTRODUCTION

Given the increasing number of applications require high-speed transmission without increasing the bandwidth of the transmission channel, this is the reason for the use of high order constellations. The Quadrature Amplitude Modulation (QAM), is highly recommended as a high order constellation. However, communication systems using QAM require a high signal to noise ratio.

ISSN: 0128-7680
e-ISSN: 2231-8526

To overcome this disadvantage, it is interesting to combine high error correction codes such as LDPC codes with QAM.

LDPC codes are linear block codes, can approach the Shannon limit (Shannon, 1948). They were proposed by Gallager in 1962 (Gallager, 1962 & Gallager, 1963) and rediscovered by Mackay in 1995 (MacKay, 1999 & Davey et al., 2002). Their decoding is done according to the principle of iterative decoding. One class of algorithms used to decode LDPC codes is commonly known message propagation algorithms (Johnson, 2010).

Message propagation algorithms are also known as iterative decoding algorithms. The first practical iterative decoding algorithm is the Sum-Product algorithm (SPA) (Gallager, 1963), also known as the belief propagation algorithm is an optimal iterative decoding algorithm but with a high computational complexity. Several algorithms have been proposed to reduce the complexity of the SPA. The messages exchanged in the SPA and its versions can be measured by the APP or the LLR depending on the type of algorithm. Therefore, the message at the decoder input must be using the same calculation. The calculation of the input messages is done to the channel output depends on the considered constellation. However, the number of operations performed by the QAM, to do the calculation of soft decisions, introduced to the decoder, increases with the constellation order. The calculation of soft decisions can be calculated using the LLR or the APP depending on the type of considered algorithm.

Several algorithms have been introduced in order to simplify the exact calculation of the LLR. The pragmatic algorithm, introduced in, attempts to simplify the calculation assuming that the likelihood values are Gaussian variables (LeGoff et al., 1994; LeGoff, 2000). The max-log-MAP algorithm is the most popular simplifying the exact algorithm (Liu et al., 2015). However, simplifications of LLR are used only for decoders based on LLR. While for decoders based on APP, we need to calculate the APP, and the latter is complex.

This work introduces a method to simplify the APP calculation. It is programmed to adapt as perfectly as possible the transmission system to the channel type. This method leads to simplify the implementation of the transmission system. In this paper, we restrict our description of combining binary LDPC code, decoded by FFT-SPA (Barnault et al., 2003 & Carrasco et al., 2008) that uses APP calculations, with square Gary-QAM constellations, over Gaussian and Rayleigh Channels. The rest of the paper is organized as follows: In section 2, a system combining a QAM and an LDPC code is presented, In section 3, the exact APP computation is given. The simplified APP computation, under Gaussian and Rayleigh channels, for square QAM constellations is presented in Section 4. Finally, the simulation results and concluding remarks are given in Sections 5 and 6, respectively.

System Combining a QAM and An LDPC Code

2^m-QAM uses a set of 2^m signals of duration T transmit m symbols $\{u_{n,i}\}$, $i \in \{1, \ldots, m\}$. From a theoretical view point, the modulation operation at the instant nT is to represent, in a two-dimensional space, the 2^m signals by a set of 2^m points called constellation, and to correspond the set of m symbols to each point of the constellation identified by its abscissa a_n and ordinate b_n.

So each set of m binary symbols is associated, at each time nT, to a couple of symbols (a_n, b_n). After passing through the transmission channel, the observation relating to the couple (a_n, b_n) is represented by a couple (a'_n, b'_n). The transmitted symbols are better follow a Gray mapping, it allows to affirm that there is usually only one erroneous symbol (Tan et al., 2014). The simplest diagram of a digital transmission system as part of the association of an LDPC code and a 2^m-QAM, is given in Figure 1.

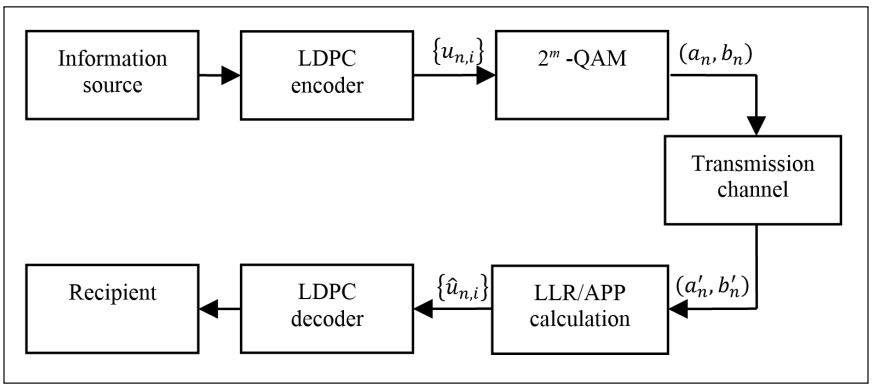


Figure 1. Diagram of a digital transmission system

At the reception, we treat couples (a'_n, b'_n) representative of the couple (a_n, b_n) to extract m samples $\{\hat{u}_{n,i}\}$, $i \in \{1, \ldots, m\}$ each representative of a binary symbol $u_{n,i}$ associated with the same signal during the modulation. Whatever the constellation with 2^m states used, for each couple (a'_n, b'_n) received at instant nT, the sample $\hat{u}_{n,i}$ is obtained using two relationships, $LLR(u_{n,i})$ (Log-Likelihood Ratio) or $APP(u_{n,i})$ (A Posteriori Probability):

$APP(u_{n,i})$, $i \in \{1, \ldots, m\}$, is calculated as follows (Moon, 2005):

$$APP(u_{n,i} = 0) = \frac{Pr\{(a'_n, b'_n) / u_{n,i}=0\}}{Pr\{(a'_n, b'_n) / u_{n,i}=0\} + Pr\{(a'_n, b'_n) / u_{n,i}=1\}} \quad (1.a)$$

$$APP(u_{n,i} = 1) = 1 - APP(u_{n,i} = 0) \quad (1.b)$$

Where $\Pr\{(a'_n, b'_n) / u_{n,i} - w\}$ is the probability that the available couple is (a'_n, b'_n); knowing the binary symbol $u_{n,i}$ is equal to w.

$LLR(u_{n,i})$, $i \in \{1, \ldots, m\}$, is calculated as follows (Alam et al., 2009):

$$LLR(u_{n,i}) = log\left[\frac{Pr\{(a'_n,b'_n) / u_{n,i}=1\}}{Pr\{(a'_n,b'_n) / u_{n,i}=0\}}\right] \quad (2)$$

Equation (2) is the exact calculation of LLR, it is the optimal calculation that represents the log-MAP algorithm (Maximum A Posteriori) (Tosato et al., 2002, Hyun et al., 2005; Wang et al., 2011). However, it involves complicated operations. Several algorithms have been introduced in order to simplify the exact calculation of the LLR.

In this work, we propose a method for applying the simplified calculation of LLR and adapt them to the LDPC decoder based on the APP, and even to simplify the APP calculation. This provided to insert an additional module to make the conversion LLR to APP, as shown in figure 2. The simplified algorithm of the APP got on the Gaussian channel can be reused efficiently on a Rayleigh channel (Figure 2), this provided insert an additional operation to accommodate, each time nT, the channel attenuation a_n.

In Figure 2, the block of the simplified calculation of the LLR is identical for all decoding algorithms, whatever the decoder input: LLR or APP. Only the modulus of the conversion of APP to LLR depends on the decoder input. His presence in the figure 2, is justified by the need to apply the simplified calculation of the LLR to LDPC decoder based on APP. Their presence is not essential in the case of an LDPC decoder based on LLR. Indeed, it is easy to change a decoding algorithm based on LLR to an algorithm based on APP, keeping unchanged the simplified calculation of LLR.

Figure 2 shows that the proposed method can simplify the implementation of the system. In what follows, we will show that this method simplifies the exact calculation of APP. The method will be explained for a QAM with 2^{2p} states, $p \in \mathbb{N}$.

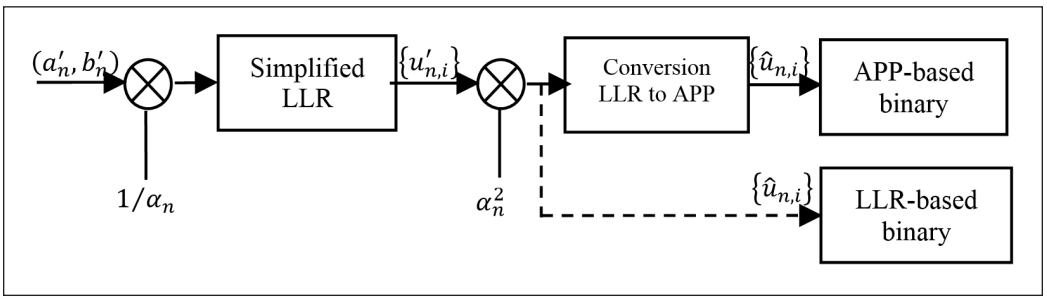


Figure 2. Principle of simplified APP calculation

Exact Calculation of APP

2^{2p}-QAM, uses a square constellation, has the particularity to be reduced to two Amplitude Modulations with 2^p states independently acting on two carriers in phase and quadrature (Ghaffar et al., 2010). According to the above property (the case of a square constellation), p expressions in phase, obtained from the equation (2) are consequently the following:

$$APP(u_{n,i}=0)=\frac{Pr\{a_n'/u_{n,i}=0\}}{Pr\{a_n'/u_{n,i}=0\}+Pr\{a_n'/u_{n,i}=1\}},\ i\in\{1,\dots,p\} \quad (3.a)$$

$$APP(u_{n,i}=1)=1-APP(u_{n,i}=0),\ i\in\{1,\dots,p\} \quad (3.b)$$

And p expressions in quadrature, obtained from the equation (2), are the following:

$$APP(u_{n,i}=0)=\frac{Pr\{b_n'/u_{n,i}=0\}}{Pr\{b_n'/u_{n,i}=0\}+Pr\{b_n'/u_{n,i}=1\}},\ i\in\{p+1,\dots,2p\} \quad (4.a)$$

$$APP(u_{n,i}=1)=1-APP(u_{n,i}=0),\ i\in\{p+1,\dots,2p\} \quad (4.b)$$

For a Gaussian transmission channel, with the noise variance σ^2, the p relations in phase eventually lead to the following expressions:

$$APP(u_{n,i}=0)=\frac{\sum_{j=1}^{2^{p-1}}exp\left\{-\frac{1}{2\sigma^2}\left(a_n'-a_{i,j}^0\right)^2\right\}}{\sum_{j=1}^{2^{p-1}}exp\left\{-\frac{1}{2\sigma^2}\left(a_n'-a_{i,j}^0\right)^2\right\}+\sum_{j=1}^{2^{p-1}}exp\left\{-\frac{1}{2\sigma^2}\left(a_n'-a_{i,j}^1\right)^2\right\}},$$
$$i\in\{1,\dots,p\} \quad (5.a)$$

$$APP(u_{n,i}=1)=1-APP(u_{n,i}=0),\ i\in\{1,\dots,p\} \quad (5.b)$$

Where $a_{i,j}^k$ are possible values of the symbol a_n when the symbol $u_{n,i}$ to be transmitted has the value k (k = 0 or 1).

Similarly, for a Gaussian channel, the p relations in the quadrature path eventually lead to the following expressions:

$$APP(u_{n,i}=0)=\frac{\sum_{j=1}^{2^{p-1}}exp\left\{-\frac{1}{2\sigma^2}\left(b_n'-b_{i,j}^0\right)^2\right\}}{\sum_{j=1}^{2^{p-1}}exp\left\{-\frac{1}{2\sigma^2}\left(b_n'-b_{i,j}^0\right)^2\right\}+\sum_{j=1}^{2^{p-1}}exp\left\{-\frac{1}{2\sigma^2}\left(b_n'-b_{i,j}^1\right)^2\right\}},$$
$$i\in\{p+1,\dots,2p\} \quad (6.a)$$

$$APP(u_{n,i}=1)=1-APP(u_{n,i}=0),\ i\in\{p+1,\dots,2p\} \quad (6.b)$$

Where $b_{i,j}^k$ are possible values of the symbol b_n when the symbol $u_{n,i}$ to be transmitted has the value k (k = 0 or 1).

Equations (5) and (6), representing the exact calculation of APP, reflects full information for all possible QAM symbols. Therefore, it is necessary a large number of calculations to calculate the exact APP in the case of a high order constellation such as 16-QAM, 64-QAM, 256-QAM. Thereafter, we will simplify the calculation, using the proposed method, when the channel is Gaussian. Then, we will show that simplified equations obtained on the Gaussian channel can be effectively reused on a Rayleigh channel.

Simplified Calculation of App

Gaussian Channel. Following the figure (3), simplification of equations (5) and (6) is obtained after the simplified calculation of LLR then converting the LLR to the APP. As the transmission channel is Gaussian, and the modulation uses a square constellation, the calculation of LLR obtained from equation (2) are well approached, using two algorithms max-log-MAP and pragmatic.

Max-log-MAP Algorithm. The Max-log-MAP algorithm, introduced in (Liu et al., 2015) shows that the p relations in the phase and p relations in the quadrature are given respectively by equation (7.a) and equation (7.b):

$$LLR(u_{n,i}) = \frac{\left(\min_{j\in\{1,\dots,2^{p-1}\}}\left(a'_n - a^0_{i,j}\right)\right)^2 - \left(\min_{j\in\{1,\dots,2^{p-1}\}}\left(a'_n - a^1_{i,j}\right)\right)^2}{2\sigma^2}, \; i \in \{1,\dots,p\} \quad (7.a)$$

$$LLR(u_{n,i}) = \frac{\left(\min_{j\in\{1,\dots,2^{p-1}\}}\left(b'_n - b^0_{i,j}\right)\right)^2 - \left(\min_{j\in\{1,\dots,2^{p-1}\}}\left(b - b^1_{i,j}\right)\right)^2}{2\sigma^2},$$

$$i \in \{p+1,\dots,2\,p\} \quad (7.b)$$

Pragmatic Algorithm. The pragmatic algorithm introduced in (LeGoff et al., 1994) shows that the p relations in the phase and p relations in the quadrature multiplied by $\sigma^2 / 2$, are given respectively by equation (8.a) and equation (8.b):

$$\begin{aligned} LLR(u_{n,1}) &= a'_n \\ LLR(u_{n,2}) &= -\left|LLR(u_{n,1})\right| + 2^{p-1} \\ LLR(u_{n,i}) &= -\left|LLR(u_{n,i-1})\right| + 2^{p-i+1} \\ LLR(u_{n,p}) &= -\left|LLR(u_{n,p-1})\right| + 2 \end{aligned} \quad (8.a)$$

And

$$\begin{aligned} LLR(u_{n,p+1}) &= b'_n \\ LLR(u_{n,p+2}) &= -\left|LLR(u_{n,p+1})\right| + 2^{p-1} \end{aligned}$$

$$LLR(u_{n,p+i}) = -|LLR(u_{n,p+i-1})| + 2^{p-i+1}$$
$$LLR(u_{n,2p}) = -|LLR(u_{n,2p-1})| + 2 \quad (8.b)$$

For a good approximation, we multiply equations (8) with a constant factor $f = 2 / \sigma^2$, we get:

$$LLR(u_{n,1}) = f \times a'_n$$
$$LLR(u_{n,2}) = f \times (-|LLR(u_{n,1})| + 2^{p-1})$$
$$LLR(u_{n,i}) = f \times (-|LLR(u_{n,i-1})| + 2^{p-i+1}) \quad (9.a)$$
$$LLR(u_{n,p}) = f \times (-|LLR(u_{n,p-1})| + 2)$$

And

$$LLR(u_{n,p+1}) = f \times b'_n$$
$$LLR(u_{n,p+2}) = f \times (-|LLR(u_{n,p+1})| + 2^{p-1})$$
$$LLR(u_{n,p+i}) = f \times (-|LLR(u_{n,p+i-1})| + 2^{p-i+1}) \quad (9.b)$$
$$LLR(u_{n,2p}) = f \times (-|LLR(u_{n,2p-1})| + 2)$$

The derivation of APP from the simplified LLR (Lee et al., 2005), leads to the following simplified equations:

$$APP(u_{n,i} = 0) = \frac{1}{1+exp(LLR(u_{n,i}))}, \; i \in \{1, \dots, 2p\} \quad (10.a)$$

$$APP(u_{n,i} = 1) = 1 - APP(u_{n,i} = 0), \; \{1, \dots, 2p\} \quad (10.b)$$

So it is remarkable that the simplified calculation of the APP, equations (10), are less number of calculations that the exact calculation of the APP, equations (5) and (6). Illustratively, consider for example the case of 16-QAM (p = 2), the functions relating the inputs and outputs, using both algorithms, pragmatic and max-log-MAP, are shown in figures 4. In each of Figures 3, there is shown the function whose form is given by equations (5) and (6) and the corresponding simplified function, equations (10), this for a signal to noise equals to 4 dB.

For bits $u_{k,3}$ and $u_{k,4}$ we have the same results as presented in figure 3.a and figure 3.b respectively. In all cases, these curves are used to verify that, despite their simplicity, the approximations performed by equations (10) are excellent, and they can therefore be used advantageously in place of expressions (5) and (6).

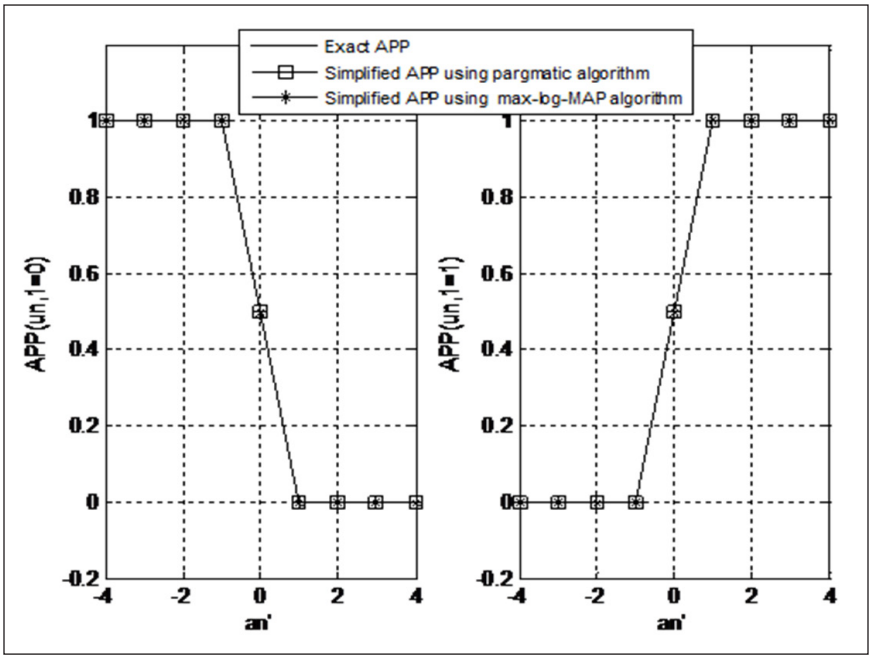


Figure 3a. Exact and simplified APP corresponding to bit $u_{n,1}$ for a 16-QAM constellation, at E_b/N_0= 4dB

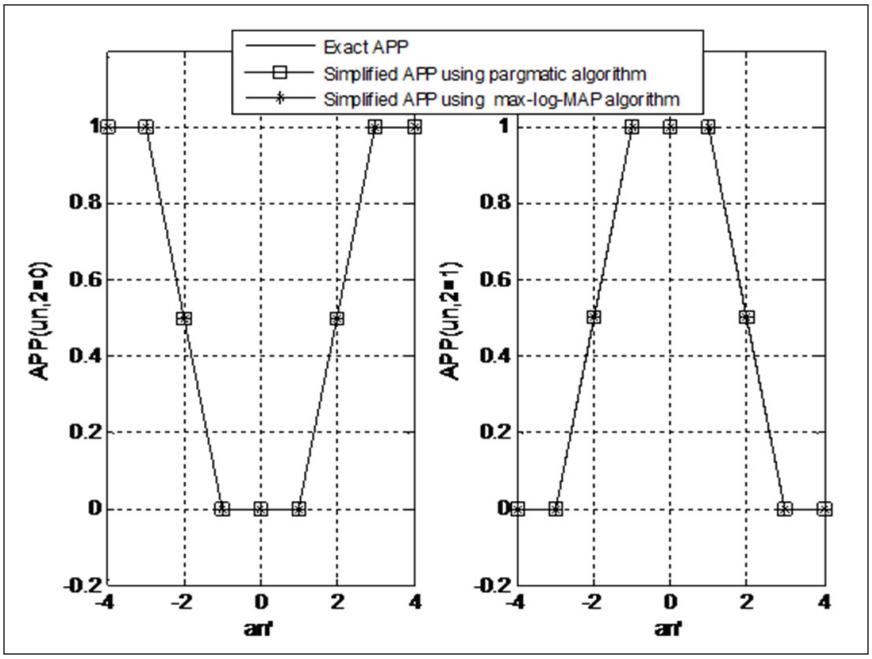


Figure 3b. Exact and simplified APP corresponding to bit $u_{n,2}$ for a 16-QAM constellation, at E_b/N_0= 4dB

Rayleigh Channel. In this section, it is assumed that the transmission channel is a Rayleigh channel. In addition, it is assumed that the attenuation of the channel a_n, at time nT, is known perfectly by the receiver (Zehevi, 1992). We will show that it is not necessary to repeat the whole argument made about the Gaussian channel to get algorithms that allow to implement the simplifications made on a Rayleigh channel (LeGoff, 1995). First, recall that the channel output at time nT, the information available relative to the couple (a_n, b_n) is such that:

$$a'_n = \alpha_n a_n + z_n \quad (11.a)$$

$$b'_n = \alpha_n b_n + z_n \quad (11.b)$$

Where z_n is a Gaussian noise, centered, with variance σ^2 and a_n variable characterizes the attenuation of the transmitted signal. As the variable a_n at time nT is known, it is possible to divide the two samples a'_n and b'_n available at the output of the channel by a_n (LeGoff, 2000). Samples a''_n and b''_n thus obtained are expressed in the form:

$$a''_n = \frac{a'_n}{\alpha_n} = a_n + z'_n \quad (12.a)$$

$$b''_n = \frac{b'_n}{\alpha_n} = b_n + z'_n \quad (12.b)$$

Where z'_n is a Gaussian noise, centered, with variance σ^2_n equals to σ^2 / a^2. Since the samples a''_n and b''_n are modeled by Gaussian variables, it is possible to apply directly on the samples a''_n and b''_n available, simplified algorithms of LLR strictly identical to those used when the transmission channel is Gaussian, and this irrespective of the modulation used. In order to apply best these results obtained with a modulation to a proposed system in this work, so it is necessary to multiply by a^2_n the samples $u'_{n,i}$ - *LLR* $(u_{n,i})$, before the conversion (Figure 2).

RESULTS

In this section, we will show the effect of simplifying calculation of the APP on the performance of binary LDPC code of rate equals to 1/2 and a parity check matrix of size 128 × 256, and a decoding algorithm using the FFT-SPA where the number of iterations is four. The transmission chain for which we evaluated the Bit Error Rate (BER) after decoding is one that has been shown in Figure 1, for a binary LDPC code, attached to two square constellations: 16-QAM and 64-QAM, using Gray mapping, and two transmission channels: Gaussian and Rayleigh.

Figure 4 shows performance comparisons, on a Gaussian channel, between a binary LDPC code using the exact calculation of the APP, equations (5) and (6), a binary LDPC code using the simplified calculation (equations (10)) by applying the pragmatic algorithm (equations (9)) and a binary LDPC code using the simplified calculation (equations (10)) by applying the max-log algorithm (equation (7)).

In the figure 4 we can see that the simplified calculation of the APP for 16-QAM,on a Gaussian channel, using a max-log-MAP algorithm and a pragmatic algorithm, has no effect on the performance of a binary LDPC code. While, with 64-QAM we have a few performance degradation when we use a pragmatic algorithm to simplify the APP calculation.

In order to study the influence of the simplified calculation on the performance of a binary LDPC code on a Rayleigh channel, the same performance comparison of Figure 4 are performed on a Rayleigh channel, in Figure 5.

Seen the results on a Rayleigh channel, there is a small performance degradation of a binary LDPC code using the pragmatic and max-log-MAP algorithms to simplify the APP.

Figure 6 presents the same performance comparison of Figures 4 and 5, using 16-QAM, on a Gaussian and Rayleigh channels, where the number of iterations is six. As illustrated in figure 6, the proposed simplified calculations are also good for binary LDPC codes with six iterations.

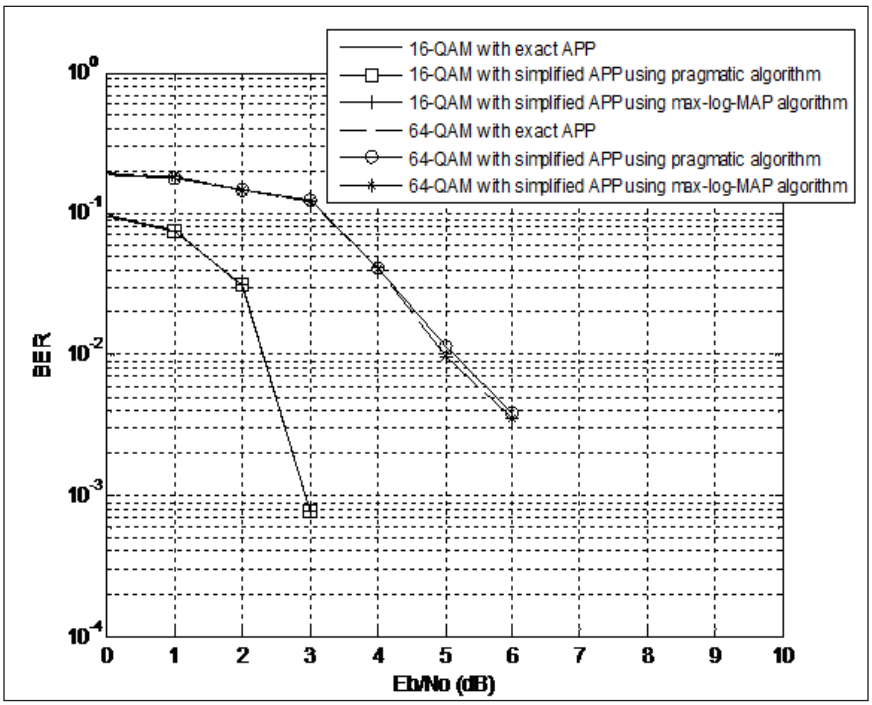


Figure 4. Performance comparisons, under Gaussian channel, of binary LDPC code decoded by FFT-SPA using exact APP and simplified APP algorithms of 16-QAM and 64-QAM constellations

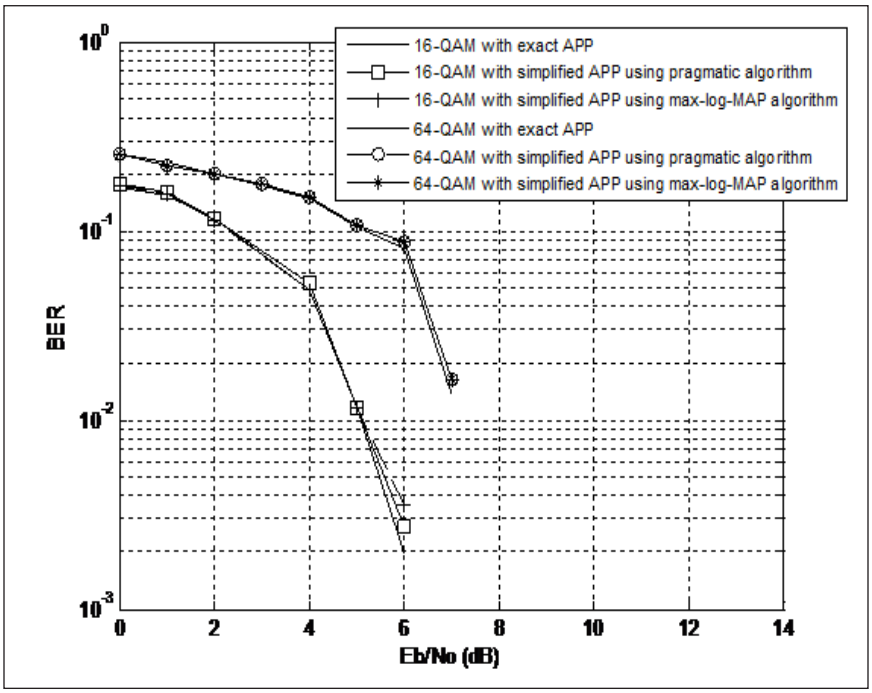


Figure 5. Performance comparisons, under Rayleigh channel, of binary LDPC code decoded by FFT-SPA using exact APP and simplified APP algorithms of 16-QAM and 64-QAM constellations

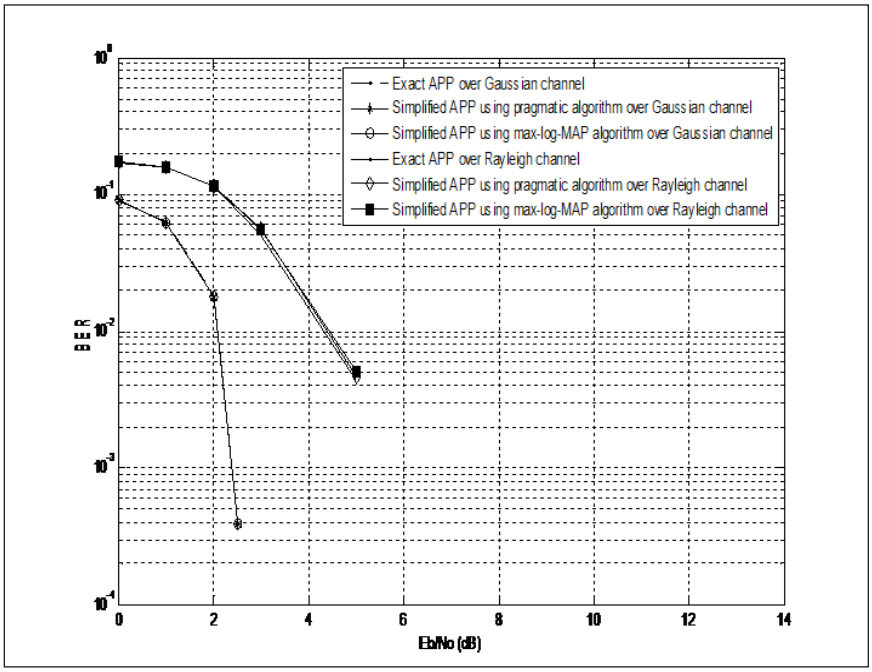


Figure 6. Performance comparisons, under Gaussian and Rayleigh channels, of binary LDPC code decoded by FFT-SPA using exact APP and simplified APP algorithms of 16-QAM with six iterations

Therefore, the simulation results presented in Figures 4, 5 and 6 show that the binary LDPC code using the proposed simplified computations can be used effectively in a system with high spectral efficiency, i. e. with high order constellations.

CONCLUSION

In this work, we used the simplified calculation of the LLR to facilitate the calculation of the APP for binary LDPC codes. The proposed method for making these simplification, puts a system combining a binary LDPC code and a high order constellation simple to implement. It is programmed to adapt as perfectly as possible the system to the type of channel in question. Also, it ensures an efficient decoding regardless of the channel in type. Since LDPC codes are selected as candidate for the 5th generation wireless communications (5G), it is interesting to use the proposed simplifications.

REFERENCES

Alam, M. Z., Islam, C. S., & Sobhan, M. A. (2009). Using Log Likelihood Relation for BER Analysis of QAM in Space Diversity. *Journal of Communications, 4*(6), 371-379.

Barnault, L., & Declercq, D. (2003). Fast Decoding Algorithm for LDPC over GF(2*q)*. In *Proceeding IEEE Information Theory Workshop* (pp. 70-73). Paris, France.

Carrasco, R. A., & Johnston, M. (2008). *Non-Binary Error Control Coding for Wireless Communication And Data Storage*. Singapore: Wiley & Sons (Asia) Pte Ltd.

Davey, M., & MacKay, D. (2002). Low-Density Parity Check Codes over GF(*q*). *IEEE Communications Letters, 2*(6), 165–167.

Gallager, R. (1962). Low-Density Parity-Check Codes. *IEEE Transaction Information Theory, 8*(1), 21–28.

Gallager, R. G. (1963). *Low-density parity-check codes* (Ph.D. dissertation). Department of Electrical Engineering, M.I.T., Cambridge, Mass.

Ghaffar, R., & Knopp, R. (2010). Low Complexity Metrics for BICM SISO and MIMO Systems. In *2010 IEEE 71st Vehicular Technology Conference (VTC 2010-Spring)* (pp. 1–6). Taipei, Taiwan.

Hyun, K., & Yoon, D. (2005). Bit Metric Generation for Gray Coded QAM Signals. *IEE Proceedings-Communications, 152*(6), 1134–1138.

Johnson, S. (2010). *Iterative Error Correction Turbo, Low-Density Parity-Check And Repeat-Accumulate Codes*. United States of America: Cambridge University Press.

Lee, S. H., Seok, J. A., & Joo, E. K. (2005). Serially Concatenated LDPC and Turbo Code with Global Iteration. In *Proceedings of the 18th International Conference on Systems Engineering (ISCEng'05)* (pp. 201-204). Las Vegas, NV, USA.

LeGoff, S. (1995). *Les Turbo-Codes et leurs Applications aux Transmissions à Forte Efficacité Spectrale*. PhD PhD Thesis. University of Western Brittany, Brest.

LeGoff, S. (2000). Bit-Interleaved Turbo-Coded Modulations for Mobile Communications. In *European Signal Processing Conference (EUSIPCO)* (pp. 1-4). Finlande.

LeGoff, S., Berrou, C., & Glavieux, A. (1994). Turbo-Codes and High Spectral Efficiency Modulations. In *IEEE International Conference on Communications (ICC'94)* (pp. 645-649). New Orleans.

Liu, X., & Kosakowski, M. (2015). Max-Log-MAP Soft Demapper with Logarithmic Complexity for M-PAM Signals. *IEEE Signal Processing Letters, 22*(1), 50-53.

MacKay, D. (1999). Good Error-Correcting Codes Based on Very Sparse Matrices. *IEEE Transaction Information Theory, 45*(2), 399–431.

Moon, T. K. (2005). *Error Correction Coding Mathematical Methods and Algorithms*. Hoboken: John Wiley & Sons, Inc.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*, 379-423.

Tan, J., Wang, Q., Qian, C., Wang, Z., Chen, S., & Hanzo, L. (2014). A Reduced-Complexity Demapping Algorithm for BICM-ID Systems. *IEEE Transactions on Vehicular Technology, 64*(9), 4350-4356.

Tosato, F., & Bisaglia, P. (2002). Simplified Soft-Output Demapper for Binary Interleaved COFDM with Application to HIPERLAN/2. In *IEEE International Conference Communication* (Vol. 2, pp. 664–668). New York, USA.

Wang, L., Xu, D., & Zhang, X. (2011). Low Complexity Bit Metric Generation for PAM Signals Based on Non-Linear Function. *Electronics Letters, 47*(17), 966–967.

Zehevi, E. (1992). 8-PSK trellis codes for a Raleigh Channel. *IEEE Transaction on Communications, 40*(5), 873-884.

Pertanika J. Sci. & Technol. 26 (4): 1765 - 1776 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

An Intelligent Reliability Assessment technique for Bipolar Junction Transistor using Artificial Intelligence Techniques

Cherry Bhargava* and Manisha Handa

Department of Electronics & Communication Engineering, Lovely Professional University, Phagwara-144411, India

ABSTRACT

The need for high speed, low cost and smaller area has increased the integration of electronic devices. As the number of components increases, the reliability of system becomes a major challenge. The bipolar junction transistor is an immensely used passive component in the various electronics industry. Reliability and failure prediction are the major constraints for the estimation of the residual life of the component. In this paper, Artificial intelligence techniques are employed on bipolar junction transistor which provides knowledge of failure mechanism of a component in actual operating conditions such that if it deviates from the actual output, a preventive measure to be taken before serious failure occurs. The end of life has been explored using the design of experiments approach. After calculating lifetime, an expert system has been modeled which predicts the sudden crash of transistor before it actual fails, using various statistical and analytical techniques. The comparison of accuracy has been conducted on all techniques of artificial intelligence and statistical method. The comparison shows that ANFIS is the most accurate technique with an accuracy of 96.65%. A graphical user interface is created which indicates the failure of bipolar junction transistor at various level of inputs.

Keywords: Artificial intelligence, design of experiments, Regression analysis, reliability prediction, Taguchi method

ARTICLE INFO

Article history:
Received: 03 November 2017
Accepted: 12 June 2018
Published: 24 October 2018

E-mail addresses:
cherry.bhargava@gmail.com/cherry_bhargav@yahoo.co.in (Cherry Bhargava)
manisha.handa@yahoo.co.in (Manisha Handa)
* Corresponding author

INTRODUCTION

Reliability assessment is the degree which tells how reliably a particular electronic system or component will work as it is expected to, in the specific or desired duration(Varde, Tian, & Pecht, 2014). The remaining useful life relies on the failure rate of a component and on the operating

ISSN: 0128-7680
e-ISSN: 2231-8526

conditions of a device. The importance of life estimation is to evaluate the remaining useful life (RUL) of a specific component under the different stress parameters (Özel & Karpat, 2005). As an increasing number of components are integrated on to a chip, the chances of failure increase, as the different parts have their own stress factors and different working conditions. So the condition monitoring strategies are utilized which enhances the reliability of a component and a suitable move to be made before any harmful breakdown happens (Kwong & Bai, 2002). The electronic circuits need a failure estimation technique to protect the system from unavoidable failures. Residual life estimation of electronic components is an important fact these days as electronic components and devices become a great need for society. Residual life prediction is predicting the remaining useful life of a component or device based on various failure factors of any component and it also depends on the operating conditions (Al-Zubaidi, Ghani, & Haron, 2013). Various methods for predicting the life of electronic components have been developed. The life of electronic components can be predicted by creating an intelligent system for the failure analysis. The capability to predict the life of electronic components is a key to prevent the sudden costly failure and it will increase the overall performance and reliability of a system (Gokulachandran & Mohandas, 2012). So, remaining useful life prediction is an important factor for every active and passive electronic component (Mamlook, Badran, & Abdulhadi, 2009).

Bipolar Junction Transistor

A transistor is an electronic gadget that controls current or voltage flow and acts as a switch or gate for electronic signals (Jang, 1993). Transistor is made up of three layers of a semiconductor material, each capable of carrying a current (Tsividis & McAndrew, 2011). It is sandwiching one semiconductor between two other semiconductors (Ramey et al., 2013). A transistor is composed of two words transfer-resistor. Mainly, there are two types of bipolar junction transistors; PNP and NPN transistor, based on doping level, as shown in Figure 1. The standard BC547 is represented as Figure 2.

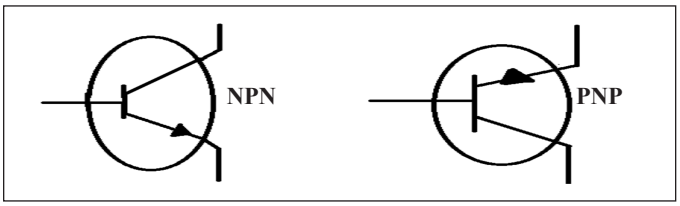


Figure 1. Symbol of bipolar junction transistor

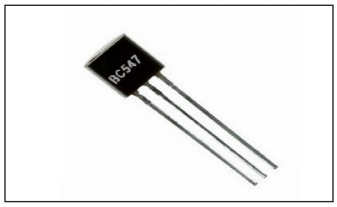


Figure 2. Bipolar junction transistor BC547

MATERIALS AND METHODS

Methods for Residual Life Prediction

Reliability assessment of electronic components is a great need of electronic society as the electronic devices are becoming integrated and high speed (Choi & Seong, 2009). The remaining useful life of electronic components, gadgets and equipment's depends on various failure factors of any component and on the operating conditions (Di Giacomo & Di Giacomo, 1997). A bipolar junction transistor using various techniques such as experimental method, statistical techniques and artificial intelligence techniques such as Regression, ANN, Fuzzy and ANFIS (Chinnam & Baruah, 2004).

Need For Failure Prediction and Life Estimation

From daily life applications to military applications and from toys to satellites, the use of electronic components is in extensive. In critical applications like aviation industry, if a component fails before its actual life, it can cost lives of many human beings. So, prediction before failure, can save the entire system and data to be lost. The user can replace faulty component with the accurate one, and system will be saved from complete shutdown (Agarwal, Paul, Zhang, & Mitra, 2007).

During manufacturing process of electronic components/devices, different tests are conducted on these components/devices to check its performance and capabilities. Then data sheets of those components have been framed, which signifies minimum and maximum range of all electrical parameters as well as environmental stress. Then components/devices are released to real market with a warranty period depending on the qualification testing (Vichare & Pecht, 2006). Before purchasing the specific component, user needs assurance about the long life, reliable and satisfactory performance throughout the operation. Moreover, quality assurance cells are becoming strict towards protection of user rights. So, the product manufacturers reacted to these constraints by offering extended warranties and guarantees. The warranty may be in the shape of lifetime i.e. the specific component will work successfully for specific duration of time, that is called lifetime. This document is generally endorsed with datasheet. The manufacturer has to replace or rectify the component free of cost, if it falls under warranty period. If the component under warranty time, could not perform the specified task, it will incur an extra cost on manufacturer to replace it or compensate the user. Replacement will not only become extra financial burden on manufacturer, but also, it will degrade the reputation of manufacturer in competitive market. The warranty service or replacement cost may vary from 2-10%, depending on sale price of that component (Murthy, 2007). The decision on warranty period is directly connected with reliability of the system or component.

The performance of next generation U.S. Army operations such as miniature driver-less ground vehicles and driver-less Arial vehicles are intensely dependent on electronics (Habtour, Drake, & Davies, 2011). Due to vibration and shocking environment, these electronic systems may experience variation in their performance. This will lead major damage to packaging and soldering of joints. Nowadays, failure rate prediction is not just a domain for the military. As electronics is used in almost every area of human life, and as a result all of these areas being "safety critical", the prediction of the lifetime of electronic modules becomes an ever-increasing necessity (Jánó & Pitică, 2011). This is especially true for aeronautics and automotive applications, where temperatures can sometimes well exceed the maximum guaranteed operating temperature for a particular component. Performance parameters as well as failure prediction, reliability and safety need to be built in during design and development and retained during operation and production of item.

Selection of Components for Life Estimation

There are many electronic components used in high-speed electronic devices nowadays. All these components have some residual life that depends on the operating conditions of a component (Debnath, Roychowdhury, & Kundu, 2016). In this paper, basic components

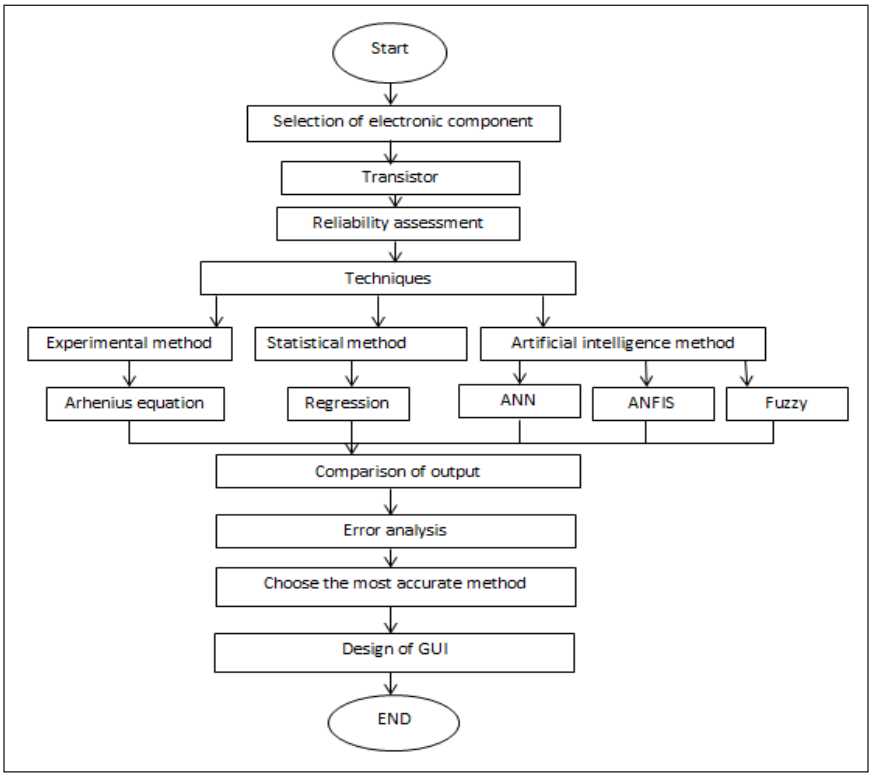


Figure 3. Methodology of the proposed work

such as a bipolar junction transistor are selected as these components are the part of many high-speed electronic devices (Novak et al., 2015). So, life estimation of these components is necessary in order to prevent the complete device from a serious breakdown and costly failures (Lu & Christou, 2017).

Methods to Exploring the Remaining Useful Life of Electronic Devices

The life estimation model is developed using various methods such as experimental method, regression and artificial intelligent model. The experimental method is the working method that predict the remaining useful life of any electronic devices(Pecht & Nash, 1994).

Experimental Method

In this method, lifetime was estimated using the ALT (Acceleration life testing). ALT is the process of testing the component under the stress factor temperature in order to find the failure rate and life of a component. This method is mainly used by big manufacturing unit where many samples or electrical units are subjected to tests (Zhao, Makis, Chen, & Li, 2018). These tests can be environmental, electrical, thermal and so on. One of the examples of such testing is Accelerated Life Testing (ALT) using temperature.

In the first step, the components were placed on the hotplate, the value of each component was measured, and the desired temperature level was set on the hot plate. The temperature was allowed to achieve the maximum rated temperature.

(a) The trial was executed for 20 hours. This time length was chosen according to various temperature ranges. Time interval must be shorter at a higher temperature as chances of failure of components are more as compared to lower temperature limit.

(b) The value of every component was measured and noted after few hours and checked on how many components got failed after few hours and calculation of the output life was done.

(c) After collection of failure data, life had been calculated using Arrhenius equation given below:

Life = $1/TDH * A_F$

Where, A_F is failure rate which is given by:

$$A_F = e^{\frac{E_a}{K}\left[\frac{1}{T_m} - \frac{1}{T_a}\right]}$$

Where, T_m is maximum temperature

T_a is applied temperature

E_a is activation energy of transistor

K is Boltzmann's constant

TDH is Total no. of devices * Hours of operation

In order to develop the life prediction model, the input parameters to be selected process

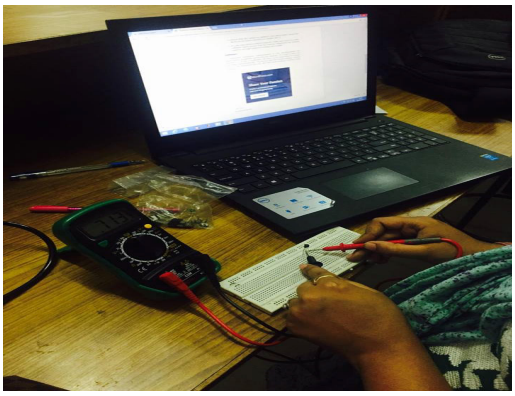


Figure 4. Experimental setup of bipolar junction transistor

Figure 5. Accelerated life testing of bipolar junction transistor

Table 1
Specification of NPN BC547

PARAMETER	VALUE
Collector-emitter voltage	45 V
Collector base voltage	50 V
Emitter-base voltage	6 V
Temperature	-55°C to 150°C
Pins	3
Collector current	100 mA

Table 2
Units, notations and limits

Process Parameters	Units	Notation	Limit Minimum	Limit Maximum
Time	Hours	T	1	49
Temperature	ºC	T	25	150

Life Estimation of Transistor Using Artificial Intelligence Techniques

The remaining useful life of bipolar junction transistor is also estimated using artificial intelligence techniques such as ANN, Fuzzy and ANFIS. Artificial intelligence techniques are the modern way of estimation as it provides the better results (Recknagel, French, Harkonen, & Yabunaka, 1997) and estimation is done in an intelligent way almost same as the human brain (Bundy, 1997).

Life Estimation using ANN. Artificial Neural Network is an analogous system of the human neural network which tries to mimic the functioning of the actual brain. Input data along with target data has been fed to the network. The system gets the train with

the specified number of the epoch. The system will train itself and reduce the error after every epoch and hence, after a specific number of epochs, the best result is achieved. The number of neurons in the input layer consists of input parameters such as temperature and time which are used to obtain the output life of the electronic component.

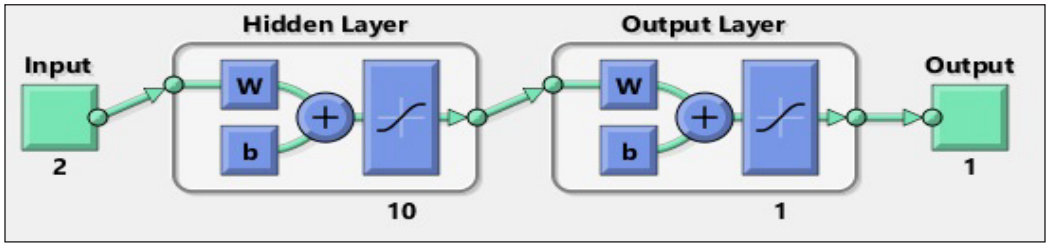


Figure 6. ANN structure (2-10-1)

Life Estimation of Bipolar Junction Transistor using Fuzzy. Fuzzy Inference System is a soft-computing technique to design intelligent model with an advantage that it is user understandable as it involves linguistic variable.

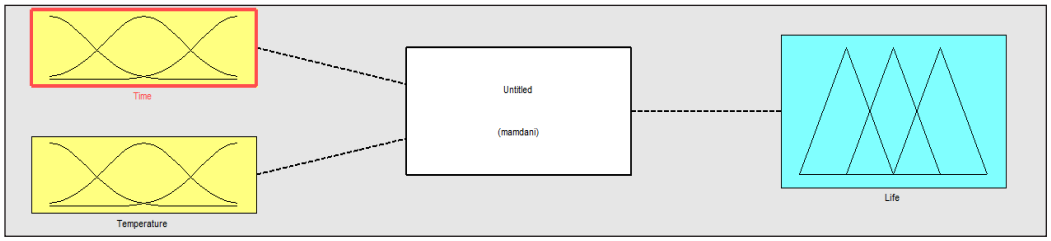


Figure 7. Fuzzy models for bipolar junction transistor

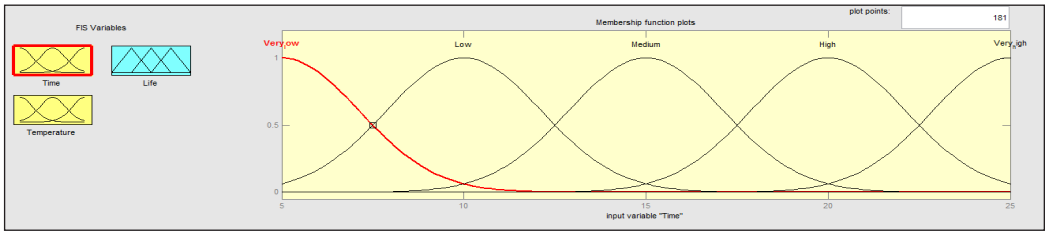


Figure 8. Gaussian membership functions for time

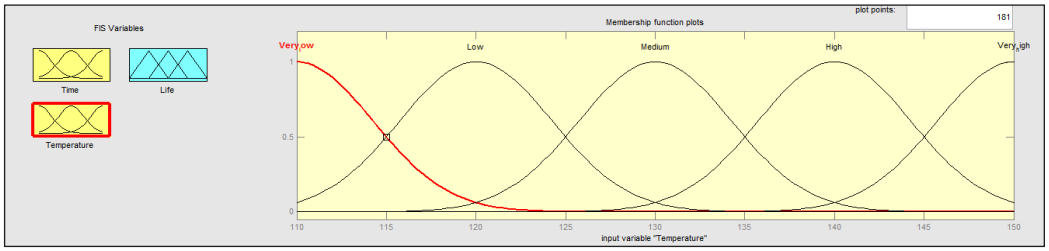


Figure 9. Gaussian membership functions for temperature

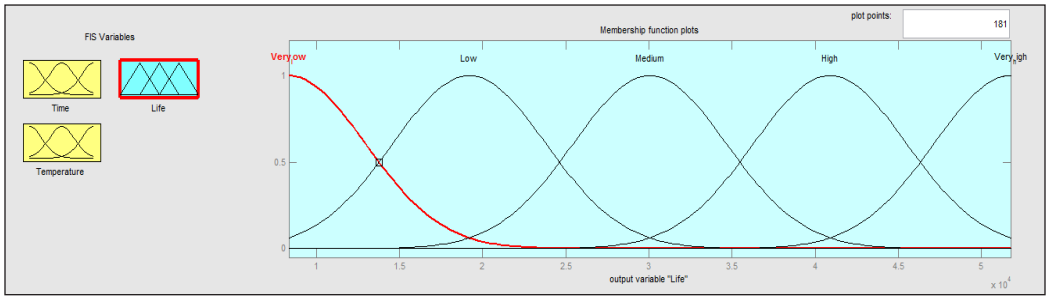


Figure 10. Gaussian membership functions for life

Life Estimation of Bipolar Junction Transistor using ANFIS. The ANFIS is an Adaptive neuro-fuzzy inference system, The ANFIS architecture shown in Figure 11, is a Sugeno fuzzy model where the final fuzzy values optimized by using the artificial neural network training (Chang & Chang, 2006).The linguistic variables very low (VL), low (L), medium (M) and high (H), very high (VH) are used for the inputs as well as for the output (Jang, 1993).

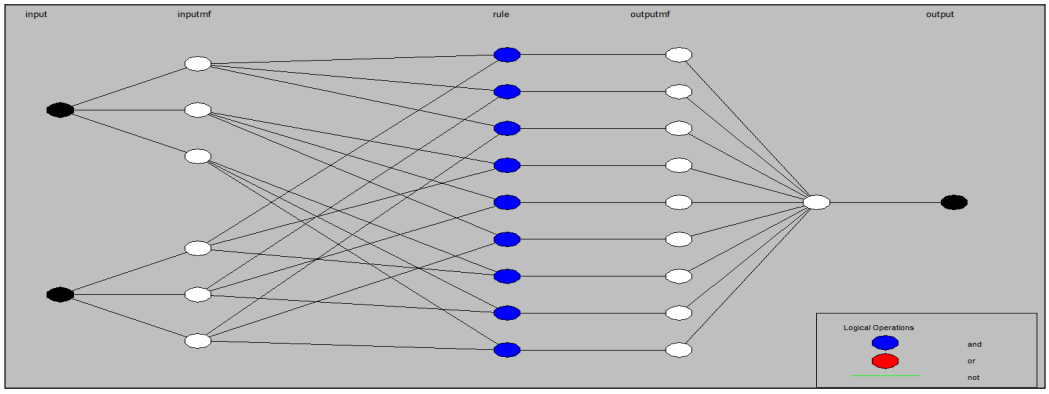


Figure 11. ANFIS structure

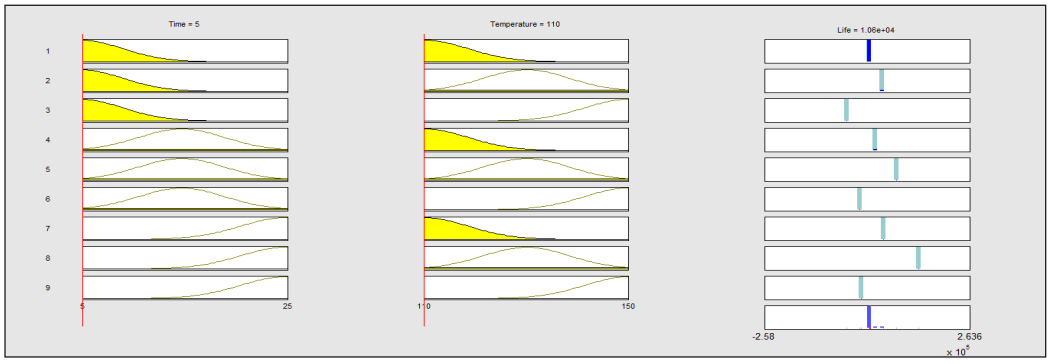


Figure 12. ANFIS rule viewer

Design of Fuzzy based Decision Support System

Designing of graphical user interface (GUI) is the last phase of this method (Aronson, Liang, & Turban, 2005).Using the GUI user can interact with the expert system to check the operating condition of the bipolar junction transistor (Ghodsypour & O'Brien, 1998). The GUI is created using the MATLAB-R2013a. The entire database including rules is designed using the fuzzy logic (Kulak, 2005). The steps to design a fuzzy based decision support system are as follows:

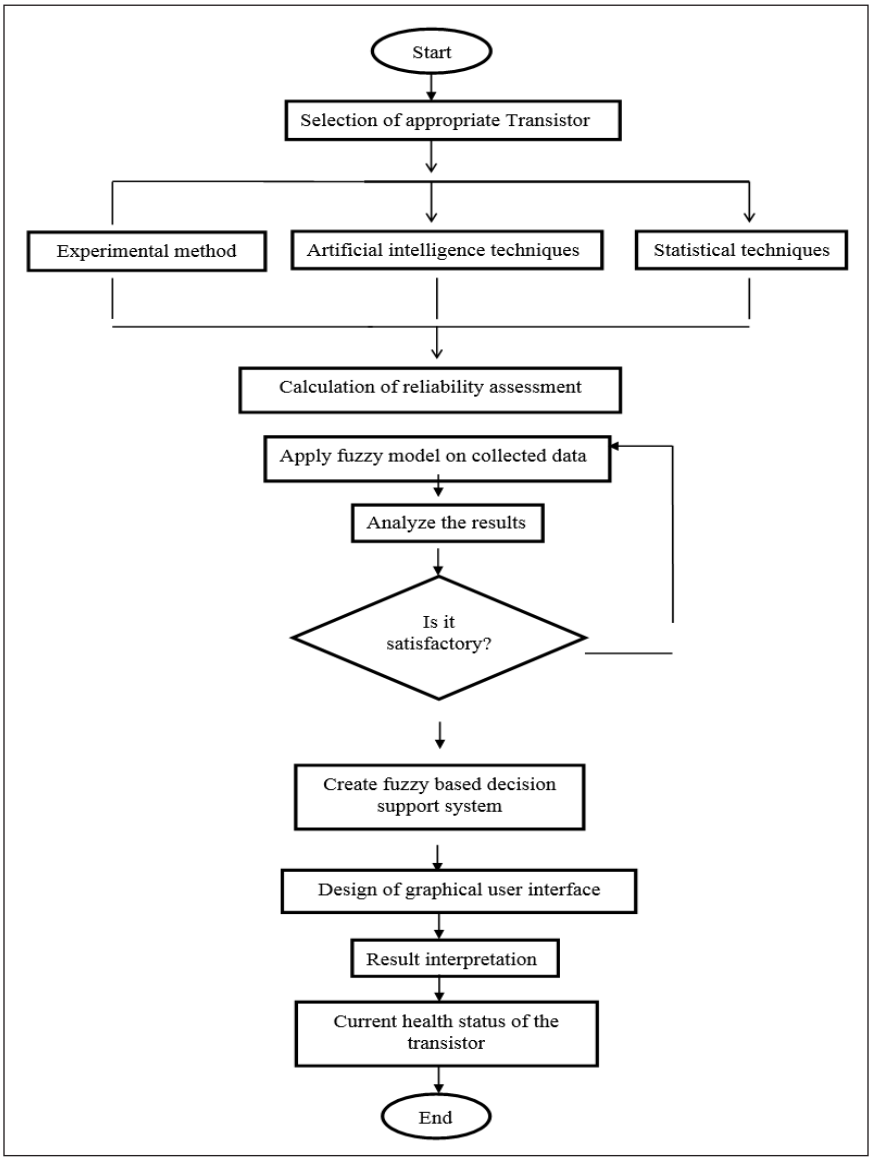


Figure 13. Flowchart of graphical user interface

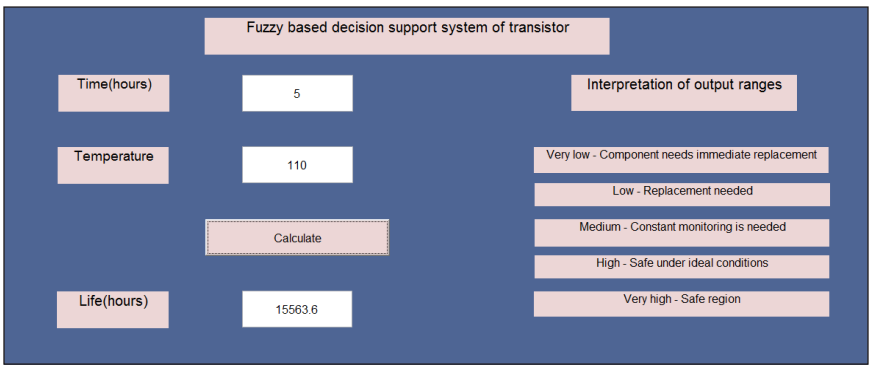


Figure 14. Graphical user interface

RESULTS AND DISCUSSION

This paper is focused particularly on life estimation of the bipolar junction transistor. Various methods have been explored to design critical parameters such as temperature. The comparison is shown as per Table 3.

Table 3
Comparison of life obtained using various methods

Techniques	Life estimation using Regression	Life estimation using ANFIS	Life estimation using ANN	Life estimation using Fuzzy
Estimated lifetime(hours)	27503	27547	26747	30330
Average error %	-15.04	-3.35	4.22	-28.56
Accuracy %	84.96	96.65	95.78	71.44

Different techniques are used for calculating the useful life of bipolar junction transistor. It is found that an Adaptive neuro inference system provides the highest accuracy i.e. 96.65%, to predict the useful life of bipolar junction transistor in comparison with other techniques.

CONCLUSION

As reliability prediction plays a significant role in the successful operation of electronics devices, it is necessary to forecast the useful life of the bipolar junction transistor for making reliable system. Experimental and mathematical analysis has been done on bipolar junction transistor under various operating conditions. It is found that the regression provides less accuracy and ANFIS has the highest accuracy to predict the useful life of electronic components. The reliability assessment of bipolar junction transistor is obtained by the experimental and artificial intelligence models to determine the relative effectiveness of

each of the developed. The accuracy of life estimated using artificial intelligence techniques such as using regression is 84.96%; ANN is 95.78%; fuzzy accuracy is 71.44% and it is observed that the adaptive neuro-fuzzy inference (ANFIS) method provides the highest rate of accuracy that is 96.65%.

REFERENCES

Agarwal, M., Paul, B. C., Zhang, M., & Mitra, S. (2007, May 6-10). *Circuit failure prediction and its application to transistor aging.* Paper presented at the 25th IEEE VLSI Test Symposium, Berkeley, USA.

Al-Zubaidi, S., Ghani, J. A., & Haron, C. H. C. (2013). Prediction of tool life in end milling of Ti-6Al-4V alloy using artificial neural network and multiple regression models. *Sains Malaysiana, 42*(12), 1735-1741.

Aronson, J. E., Liang, T. P., & Turban, E. (2005). *Decision Support Systems and Intelligent Systems.* Upper Saddle River, NK: Prentice Hall.

Bundy, A. (1997). Artificial intelligence techniques. In *Artificial Intelligence Techniques* (pp. 1-129). Berlin, Heidelberg: Springer.

Chang, F. J., & Chang, Y. T. (2006). Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources, 29*(1), 1-10.

Chinnam, R. B., & Baruah, P. (2004). A neuro-fuzzy approach for estimating mean residual life in condition-based maintenance systems. *International Journal of Materials and Product Technology, 20*(1-3), 166-179.

Choi, J. G., & Seong, P. H. (2009). Reliability of electronic components. In *Reliability and Risk Issues in Large Scale Safety-critical Digital Control Systems* (pp. 3-24). Heidelberg: Springer.

Debnath, B., Roychowdhury, P., & Kundu, R. (2016). Electronic Components (EC) Reuse and Recycling–A New Approach towards WEEE Management. *Procedia Environmental Sciences, 35*, 656-668.

Di Giacomo, G., & Di Giacomo, G. (1997). *Reliability of Electronic Packages and Semiconductor Devices*. New York: McGraw-Hill.

Ghodsypour, S. H., & O'Brien, C. (1998). A decision support system for supplier selection using an integrated analytic hierarchy process and linear programming. *International Journal of Production Economics, 56*, 199-212.

Gokulachandran, J., & Mohandas, K. (2012). Predicting remaining useful life of cutting tools with regression and ANN analysis. *International Journal of Productivity and Quality Management, 9*(4), 502-518.

Habtour, E., Drake, G. S., & Davies, C. (2011, January 24-27). *Modeling damage in large and heavy electronic components due to dynamic loading.* Paper presented at the IEEE Annual Symposium on Reliability and Maintainability Symposium (RAMS), Lake Buena Vista, USA.

Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics, 23*(3), 665-685.

Jánó, R., & Pitică, D. (2011, May 11-15). *Parameter monitoring of electronic circuits for reliability prediction and failure analysis.* Paper presented at the IEEE 34th International Spring Seminar on Electronics Technology (ISSE), Tratanska Lomnica, Slovakia.

Kulak, O. (2005). A decision support system for fuzzy multi-attribute selection of material handling equipments. *Expert Systems with Applications, 29*(2), 310-319.

Kwong, C. K., & Bai, H. (2002). A fuzzy AHP approach to the determination of importance weights of customer requirements in quality function deployment. *Journal of Intelligent Manufacturing, 13*(5), 367-377.

Lu, Y., & Christou, A. (2017). Lifetime Estimation of Insulated Gate Bipolar Transistor Modules Using Two-step Bayesian Estimation. *IEEE Transactions on Device and Materials Reliability, 17*(2), 414-421.

Mamlook, R., Badran, O., & Abdulhadi, E. (2009). A fuzzy inference model for short-term load forecasting. *Energy Policy, 37*(4), 1239-1248.

Murthy, D. (2007). Product reliability and warranty: an overview and future research. *Production, 17*(3), 426-434.

Novak, S., Parker, C., Becher, D., Liu, M., Agostinelli, M., Chahal, & M., Natarajan, S. (2015, April 19-13). *Transistor aging and reliability in 14nm tri-gate technology.* Paper presented at the IEEE International Reliability Physics Symposium (IRPS 2015), Monterey, USA.

Özel, T., & Karpat, Y. (2005). Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks. *International Journal of Machine Tools and Manufacture, 45*(4), 467-479.

Pecht, M. G., & Nash, F. (1994). Predicting the reliability of electronic equipment. *Proceedings of the IEEE, 82*(7), 992-1004.

Ramey, S., Ashutosh, A., Auth, C., Clifford, J., Hattendorf, M., Hicks, J., & St Amour, A. (2013, April 14-18). *Intrinsic transistor reliability improvements from 22nm tri-gate technology.* Paper presented at the IEEE International Reliability Physics Symposium (IRPS), Anaheim, USA.

Recknagel, F., French, M., Harkonen, P., & Yabunaka, K. I. (1997). Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling, 96*(1-3), 11-28.

Tsividis, Y., & McAndrew, C. (2011). *Operation and Modeling of the MOS Transistor*. New York, NY: Oxford University Press.

Varde, P. V., Tian, J., & Pecht, M. G. (2014, July 28-30). *Prognostics and health management based refurbishment for life extension of electronic systems.* Paper presented at the 2014 IEEE International Conference on Information and Automation (ICIA), Hailar, China.

Vichare, N. M., & Pecht, M. G. (2006). Prognostics and health management of electronics. *IEEE Transactions on Components and Packaging Technologies, 29*(1), 222-229.

Yu, Y., Zhou, T., Zhu, M., & Xu, D. (2012, December 8-10). *Fault diagnosis and life prediction of dc-link aluminum electrolytic capacitors used in three-phase ac/dc/ac converters.* Paper presented at the Second IEEE International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMCCC), Harbin, China.

Zhao, S., Makis, V., Chen, S., & Li, Y. (2018). Evaluation of reliability function and mean residual life for degrading systems subject to condition monitoring and random failure. *IEEE Transactions on Reliability, 67*(1), 13-25.

Pertanika J. Sci. & Technol. 26 (4): 1777 - 1788 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Rebuilding Hydrological Data with ANN or GA Methods: Case Study-Dez Reservoir, Western Iran

Adib, A.

Civil Engineering Department, Engineering Faculty, Shahid Chamran University of Ahvaz, Ahvaz, Iran

ABSTRACT

Access to sufficient and confident hydrometric data is necessary for water resources management. Most of the Iran's hydrometric stations do not have sufficient data. The method of producing synthetic data should use probability concepts and retains main characteristics of the data, too. In this research, synthetic hydrometric data are generated by the monthly and annual Markov chain method at the Telezang station in the upstream of the Dez River. Using the discharge of the driest day and the wettest day of each month and the generated monthly hydrometric data, the probable highest and lowest daily discharge for each month was calculated. At the end, artificial neural network was trained with a number of observed and generated hydrometric data. The results of artificial neural network were compared with a number of observed hydrometric data which were not used in training of the network. The training of artificial neural network (ANN) with the generated hydrometric data can improve results of network. For more improvement of the results of network, genetic algorithm (GA) is used in its training and optimizing its parameters. The GA method can reduce the MSE (mean of square error) by 97% that of ANN.

Keywords: Artificial neural network, Dez River, genetic algorithm, Markov chain, Telezang station

INTRODUCTION

The forecasting of drought periods is an important task for water resources management. Therefore, accessibility to sufficient and accurate climatic and hydrometric data is necessary. Because of the shortage of hydrometric stations and inaccuracy of their data in developing countries, using of methods for generation of data is essential.

Occurrence of drought is a usual and destructive phenomenon in the Middle East countries. Application of a stochastic method is appropriate for prediction of drought, wet periods and their duration.

ARTICLE INFO

Article history:
Received: 01 November 2017
Accepted: 27 July 2018
Published: 24 October 2018

E-mail addresses:
arashadib@yahoo.com

ISSN: 0128-7680
e-ISSN: 2231-8526

This method must generate synthetic flow discharge data by considering observed flow discharge data (by saving of stochastic characteristics of observed data as mean, variance and governing stochastic distribution). Markov chain method is one of the suitable methods for this purpose. A new method must be utilized for verification of the results of the Markov chain method. This new method can generate synthetic flow discharge data. The generated data with the new method should be compared to observed flow discharge data and generated flow discharge data by Markov chain method. In this regard, artificial neural network is an appropriate method and has necessary stated characteristics.

Hydrologists utilized different Markov chain methods for generating daily and monthly flow discharge in rivers. Xu et al. (2001) utilized MACP (Markov Auto- Correlation Plus) method for generation of daily, monthly and annual flow discharge in the Wupper River in Germany. They studied about spatial correlation between two adjacent stations. Also Xu et al. (2003) used MCCP (Markov Cross- Correlation Plus) method for forecasting daily stream flow of the Wupper River in Germany. They determined wet and dry periods. Aksoy (2003) utilized Markov chain method for prediction of daily stream flow. He determined wet and dry days in different watersheds. Szilagyi et al. (2006) used hybrid Markov chain method for forecasting daily stream flow of the Tisza River and its top branches (Szamos, Bodrog, & Kraszna) in Hungary. They utilized observed data from 1951 to 2000 and considered two dry and wet states. They evaluated transfer probability between two states (dry to dry, wet to dry, dry to wet and wet to wet). Sarlak et al. (2009) predicted annual stream flow of the Goksu River in Turkey based on oscillation of sea water level in North Atlantic Ocean. Stošić et al. (2012) applied Monte Carlo Markov Chain (MCMC) for prediction of discharge and velocity profile in the Exu River and the Capibaribe River. These rivers are in the northeast of Brazil.

Also, some of the hydrologists used ANN for generation of monthly flow discharges in rivers. Anmala et al. (2000) applied a feed forward ANN and a recurrent ANN for prediction of runoff in three watersheds of Kansas, USA. The inputs layer of ANN included monthly precipitation and temperature and the output layer included monthly runoff. Cigizoglu (2003) utilized combination of artificial neural network and autoregressive–moving-average (ARMA) model for forecasting of monthly stream flow of Karahacili hydrometric station on the Goksu River in Turkey (east of Mediterranean Sea). He used ~~of~~ perceptron ANN and trained it with the observed and the data generated by ARMA model. The MSE of the ANN trained with the generated data by ARMA model was less than that of the ANN trained with the observed data. Keskin and Taylan (2010) used combined McCulloch and Pitts ANN and adaptive neuro-fuzzy inference system (ANFIS) methods for prediction of monthly stream flow of the Alara River in Turkey. They concluded that ANN could predict monthly stream flow better than ANFIS. Raman and Sunilkumar (2010) applied a perceptron ANN for generation of monthly stream flow of main river of Bharathapuzha

watershed in India. They concluded that ANN could predict monthly stream flow better than ARMA model. Ochoa-Riveria et al. (2002) used hybrid artificial neural network for generation of monthly flow discharge. Results of their network were more accurate than results of ARMA model. Ahmed and Sarma (2007) applied ANN, ARMA and Thomas-Fiering models for prediction of monthly stream flow of the Pagladia River (a top branch of the Brahmaputra River) in Himalayan region of India. They observed that ANN could predict monthly stream flow more accurately than other methods. Lee and Kang (2016) used ANN for simulating daily flow discharge in the Bocheong-cheon watershed (in the centre of South Korea). Parsaie et al. (2017) applied ANFIS method in prediction of flow discharge in the compound open channel. The coefficient of determination (0.98) and root mean square error (0.029) illustrated the accuracy of ANFIS method. Young et al. (2015) combined two hybrid models and Hydrologic Engineering Center- Hydrologic Modeling System (HEC-HMS) model. They linked this combined model to genetic algorithm neural network (GANN) and ANFIS and applied it in the Laonong Creek basin in southern Taiwan. This method had a high accuracy in prediction of hourly runoff discharge. Khan et al. (2016) applied ANN for forecasting of flow discharge and water surface elevation in the Ramganga River catchment of the Ganga Basin (in India). The mean square errors (MSE) of flow discharge and water surface elevation were 0.046 and 0.012.

The biggest gap and problem of the above researches is that they use only one method which reduces the accuracy of generated data. Therefore, some of researchers combined the Markov chain and ANN methods for forecasting monthly and annual flow discharges, rather than using only one of these methods. Therefore, they covered the gap between the applications of each method separately. For example, one can refer to the following researches:

Adib and Mahmoodi (2017) utilized the Markov chain method for prediction of flow discharge in Idenak hydrometric station, located at the Marun River. This river is located in southwest of Iran. They predicted suspended sediment load using ANN and generated flow discharges by the Markov chain method. Also, they optimized the parameters of ANN with the GA and the GA could reduce the Normalized Mean Square Error (NMSE) by 20%. Rezaeianzadeh et al. (2016) used the Markov chain method and ANN for prediction of inflow discharge to the Doroodzan reservoir in south of Iran. A combination of the two methods increased the accuracy of forecasting the droughts and flow discharges. The results of this combined method are more reliable too.

In this research, a perceptron ANN was utilized for prediction of mean monthly and annual discharge in hydrometric stations. For training of this network, observed and generated data by Markov chain method were used. The GA method was used in optimization of parameters of ANN for improving ANN method.

MATERIALS AND METHODS

The Telezang Hydrometric Station of the Dez River

A hydrometric station was considered on the Dez River (Telezang station), for generating the synthetic data. This station is at the upstream of the dams constructed on the Dez River. Therefore, these dams cannot regulate discharge of fluvial flow in this station. This hydrometric station was constructed in 1955 and its elevation is 468 m above sea level. The area of its watershed is 16130 km^2. It is in 48°46'3" E and 32°49'19" N. The mean of annual flow discharge, precipitation and temperature are 249.569 CMS, 76.2 mm and 33.6°C, respectively, in the Telezang hydrometric station. The vicinity of this station is shown in Figure 1 and the stochastic parameters of different months are illustrated in Table 1. The mean and variance of monthly and annual flow discharges in this table pertain to data of several years (1977-2015).

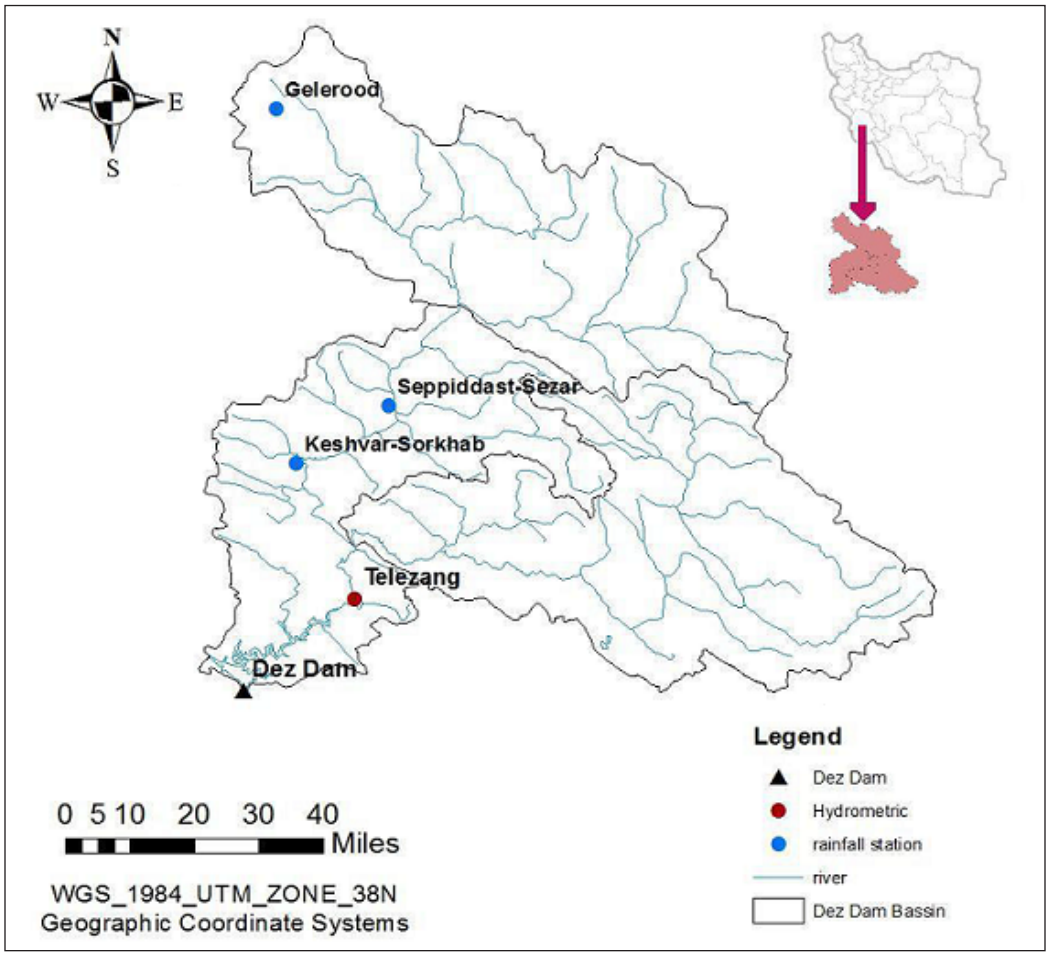


Figure 1. The region of Telezang Station- Banihabib et al. (2017)

Table 1
Stochastic parameters of the Telezang station

Month	Mean (CMS)	Variance (CMS^2)	Governing stochastic distribution
Jan	206.512	14390.96	Log PersonIII
Feb	293.9	33002.18	Log PersonIII
Mar	418.575	68039.38	Log PersonIII
Apr	580.092	77971.41	Log PersonIII
May	510.13	64629.52	Log PersonIII
Jun	276.282	14813.83	Log PersonIII
Jul	165.122	4328.768	Log PersonIII
Aug	110.712	1881.624	Log PersonIII
Sep	81.535	762.552	Log PersonIII
Oct	68.643	393.278	Log normal 2 Para
Nov	101.608	4767.981	Log PersonIII
Dec	181.712	21875.2	Log PersonIII
Year	249.569	8915.936	Log PersonIII

METHODS

In this research, monthly and annual Markov chain methods were applied.

This method uses (1) for annual Markov chain method:

$$Q_{i+1} = \overline{Q} + r_Q(1)(Q_i - \overline{Q}) + t_{i+1} S_Q (1 - r_Q(1)^2)^{1/2} \quad [1]$$

Where: Q is flow discharge (CMS), S is the standard deviation of flow discharges (CMS), r is correlation coefficient between successive flow discharges and t is the value of t-distribution.

For monthly Markov chain method (1) is converted to (2) and (3).

$$Q_{i,j+1} = \overline{Q}_{j+1} + b_j (Q_{i,j} - \overline{Q}_j) + t_{i,j+1} S_{Q_{j+1}} (1 - r_j^2)^{1/2} \quad [2]$$

$$b_j = r_j \frac{S_{j+1}}{S_j} \quad [3]$$

Where: i is index of year and j is index of month.

In this research, a series of random numbers were produced using (4):

$$P_{i+1} = \text{Mod}(aP_i + b)/c \quad [4]$$

Where: P is random number and a, b and c are constants chosen optionally by the users. The number of digits of the random number is equal to the number of digits of c.

Neural networks can be divided into two types based on their structures: feed forward networks and recurrent networks. In this research, a feed forward network is used. Application of feed forward networks is more common than of recurrent networks in water engineering.

In a feed forward network, the nodes are grouped into layers. Signals flow from the input layer through the network towards the output layer, via unidirectional connections. The nodes are connected from one layer to the next one, but not within the same layer. A multi-layer perceptron (MLP) is a feed forward network with one or more hidden layer. Given a training set of input-output data, the most common learning rule for multi-layer perceptrons is the back propagation algorithm. A neural network with such a type of learning algorithms is usually referred to as back propagation network (BPN).

The genetic algorithm method is utilized for optimizing the training step of artificial neural network. This procedure modifies parameters of artificial neural network such as momentum and learning rate. However, this method increases the time of training step, as the genetic algorithm, searches for the global optimum of fitness between results of artificial neural network and observed data. Also, the fitness between the results of this method and those of the Markov chain method is better than the fitness between the results of ordinarily trained artificial neural network and Markov chain methods.

Also mean square error (MSE) is a performance criterion in this research.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Q_{gen} - Q_{obs})^2 \qquad [5]$$

Where: n is number of data, Q_{gen} is generated streamflow discharge and Q_{obs} is observed streamflow discharge.

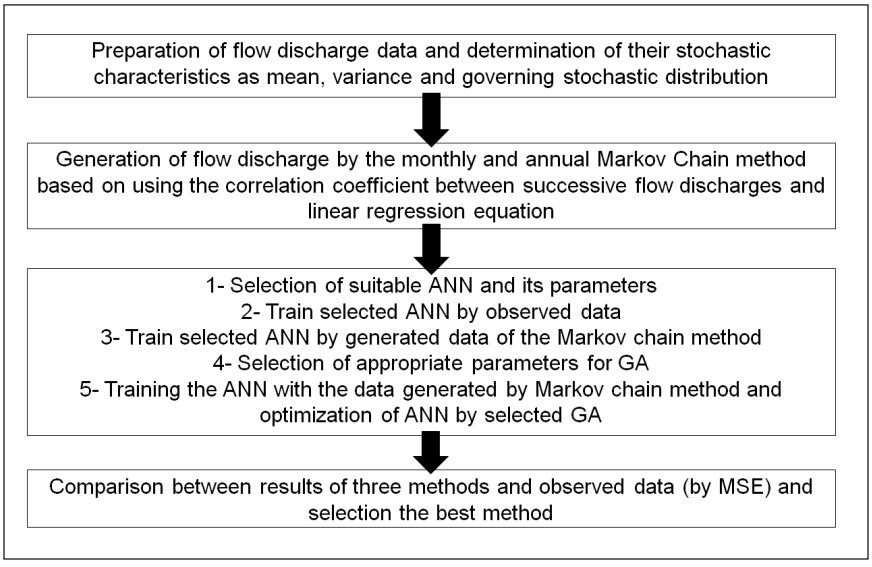


Figure 2. The Flowchart of this research methodology

RESULTS AND DISCUSSION

Results of Monthly and Annual Markov Chain Method

Eighty time series of hydrometric data were generated with the monthly and annual Markov chain method. The number of data in each time serial is equal to 50. If a data of a time series is lower than the mean of the that time series, it shows a drought period, also, if a data of a time series is greater than the mean of that time series, then it shows a wet period. In this research, 100 time series were generated with the monthly and annual Markov chain method. Each annual time serial and monthly (for 12 months) time serial has 50 data. The mean of annual and monthly generated data (50*100 data is number of generated annual data and data of each month) must be compared to annual and monthly observed data. The driest time serial, among the produced 100 time series of generated annual and monthly data, is the one with the minimum mean of its monthly flow discharge. Also, the wettest time serial is the one that with maximum mean of its monthly flow discharge. The driest and wettest annual and monthly time series should also be compared to observed annual and monthly data too.

The results of monthly and annual Markov chain method are illustrated in Table 2.

Table 2
Comparison between generated time series and observed data in the Telezang Station

Month	Difference between the mean of discharges of generated the driest time serial and the mean of observed data (%)	Difference between the mean of discharges of generated the wettest time serial and the mean of observed data (%)	Difference between the mean of discharges of generated time series and the mean of observed data (%)
Jan	-15.26	14.22339	-0.71667
Feb	-17.782	16.15243	-1.01837
Mar	-17.498	15.33035	-1.07149
Apr	-14.829	9.971349	-1.17826
May	-13.805	9.750652	-1.05659
Jun	-6.3949	4.624261	-0.74236
Jul	-4.0243	2.957207	-0.35731
Aug	-4.8369	2.801864	-0.48956
Sep	-4.1209	2.60379	-0.26246
Oct	-8.9011	7.380214	-0.7881
Nov	-17.18	29.75947	-0.872
Dec	-16.315	35.02355	0.310381
Year	-14.005	12.16097	-1.162

Results of Daily Markov Chain Method

In this research, a new method was applied for determination of the driest and wettest days in each month. The method is based on the following procedure:

The discharge of the driest day generated in a particular month is equal to the product of observed discharge of the driest day in that particular month and the ratio of the average of monthly discharges of the generated driest time series and the average of the observed data of corresponding time series.

$$Q_{Gen\ the\ driest\ day} = Q_{Obs\ the\ driest\ day} * (Q_{Avg\ of\ gen\ the\ driest\ time\ series} / Q_{Avg\ of\ obs\ data}) \quad [6]$$

Also, the discharge of the generated wettest day in each specific month is equal to product of discharge of the observed wettest day in that particular month and the ratio of the mean of discharges of the generated wettest time series and the mean of the observed data in that time series.

$$Q_{Gen\ the\ wettest\ day} = Q_{Obs\ the\ wettest\ day} * (Q_{Avg\ of\ gen\ the\ wettest\ time\ series} / Q_{Avg\ of\ obs\ data}) \quad [7]$$

The results of daily Markov chain method are illustrated in Table 3.

Table 3
Generated discharges of the driest and wettest day in the Telezang Station

Month	Generated discharge of the driest day (CMS)	Generated discharge of the wettest day (CMS)
Jan	46.61	3213.1
Feb	60.02	5216.4
Mar	84.98	3658.28
Apr	114.1291	3179.27
May	113.78	3492.27
Jun	78.44	1011.72
Jul	57.2	519.93
Aug	44.73	306.35
Sep	36.34	226.75
Oct	33.71	164.29
Nov	34.78	4130.24
Dec	49.54	6115.22

Estimation of the Longest Drought Period

The longest observed drought period in the Dez watershed was from 1957 to 1968 (11 years). The Markov chain method could generate it. The longest generated drought period is, also, equal to 11 years. The longest drought period is the longest number of consecutive

years that their annual flow discharges are less than the mean annual flow discharge of the same time series data. It means that the annual flow discharge from 1957 to 1968 was less than 249.569 CMS (mean observed annual flow discharge in the Telezang station).

Results of ANN and ANN Trained with the GA

The GA is a popular method in civil and water engineering for optimization of nonlinear problems. Training with the GA method can improve the results of ANN and reduce its MSE. This method can easily be linked to numerical models and ANN, using MATLAB toolboxes. While regular gradient-based technique can not optimize the non linear problems.

The objective function of the GA method is:

$$\text{Objective function: } Min\sum_{i=1}^{m}(\text{Output of ANNs} - \text{Desired output})^2 \quad [8]$$

In order to stop the training process of ANN, a convergence criterion must be considered. This criterion is shown in below:

$$\text{Abs (Output of ANN- Desired output)} < \text{error tolerance} \quad [9]$$

The training of ANN will be terminated if this criterion is satisfied for all of the outputs of ANN.

Three methods of training, used in this research, include:

1- Training of ANN with the observed data

2- Training of ANN with the data generated by the Markov chain method

3- Training of ANN with the data generated by the Markov chain method and optimization of the parameters of ANN with the GA method

Sixty percent of data were used for training of ANN. Also 15 percent of the data were used in validation of ANN and the remaining 25 percent were utilized in testing process of ANN. The best architecture and parameters of perceptrons ANN are selected using the trial error method. This network has one input node (the mean discharge of previous month or year), one output node (the mean of discharge of present month or year) and one hidden layer with two nodes. The momentum and learning rate of network are 0.6 and 0.1 respectively. The transfer function is assumed tangent hyperbolic. This network utilizes from back propagation for training and error tolerance of ANN is 0.01.

The GA method optimizes momentum and learning rate of network. The GA method results in a momentum and learning rate of 0.55 and 0.08, respectively. The characteristics of the applied GA in this research are:

Rate of crossover=0.8, Type of mutation= Uniform, Type of crossover= Heuristic, Selection method= Stochastic universal sampling, Number of generations= 3000, Population of each generation= 120. Also, mutation rates for different generations are:

Mutation rate=0.3 if (no of generation<700)

Mutation rate= (-0.295/1300)*(no of generation-700) +0.3 if (700<no of generation<2000)

Mutation rate=0.005 if (no of generation>2000)

On average, the training of ANN with synthetic data reduces the mean of square error (MSE) by 48%, whilst the training of ANN with synthetic data followed by optimization of ANN with the GA method further reduces the MSE by 97% overall. For example while the MSE of annual flow discharge in the training stage of a regular ANN is 198182.88 CMS^2, it is 4999.59 CMS^2 for NN+GA+MARKOV, showing a 97.48% reduction, 100-(4999.59*100/198182.88)=97.48%. The MSE of different training methods of ANN are presented in Tables 4 and 5 for training and testing stages respectively. Also in this research, a recurrent network with similar parameters and architecture to applied MLP network was used. But, the MSE of the recurrent network was almost twice the MSE of applied MLP network. Therefore MLP network was selected as superior network.

Table 4
Comparison of MSE of different training methods of ANN at the Telezang Station (training stage)

Month	MSE $(CMS)^2$		
	ANN+GA+MARKOV	ANN+ MARKOV	ANN
Jan	5364.44	98912.1	177597.03
Feb	22058.71	262275.84	447604.5
Mar	16271.18	397336.8	802419.6
Apr	39124.15	908614.8	1758405.9
May	16949.89	435803.1	886870.2
Jun	4170.12	100122	203548.2
Jul	1707.82	42817.17	84591.91
Aug	877.9	21666.73	41993.81
Sep	586.9	14789.3	19200.11
Oct	419.92	7333.62	14061.47
Nov	714.98	18190.14	33286.65
Dec	4347.4	73732.38	134681.82
Year	4999.59	102597.9	198182.88

Table 5
Comparison of MSE of Different Training Methods of ANN in the Telezang Station (testing stage)

Month	MSE $(CMS)^2$		
	ANN+GA+MARKOV	ANN+ MARKOV	ANN
Jan	6200.732	112017.6	195701.1
Feb	28060.53	293814.6	501693.5
Mar	16040.68	469974.5	968437.9

Table 5 *(continue)*

Month	MSE $(CMS)^2$		
	ANN+GA+MARKOV	ANN+ MARKOV	ANN
Apr	45594.35	958903.2	1836703
May	25781.85	571080.6	1135659
Jun	7039.602	121709.7	236454.5
Jul	2589.605	44962.59	85645.8
Aug	1211.123	22064.8	41277.78
Sep	640.8639	13046.96	24454.46
Oct	582.5766	9404.678	18058.33
Nov	720.6559	16845.31	30839.54
Dec	5274.558	70838.36	124535.5
Year	6763.516	126010.8	239793.2

CONCLUSION

In this research, it was proven that the Markov chain method could generate hydrometric data. This method can predict the duration of drought period precisely. Also, it produces both dry and wet time series. By generation of dry time series, managers and designers can prepare plans for water resource management under critical conditions and meet the water demand. In the other hand, to control the flood damages, designers can use wet time series generated by Markov chain method. The mean of generated time series data and observed data are very close. This proves the accuracy of Markov chain method of data production.

ANN trained with the data generated by Markov chain method and optimized with the GA is the best method for evaluation of correctness, (verification), of generated data. The data produced by this network have the best fitness to observed data. The MSE of this method was very low. Also, the MSE of ANN trained with the data generated by Markov chain method was less than the MSE of ANN that trained with the observed data. This proved that the variation of data generated by Markov chain method was considerably more than that of the observed data, and training ANN with the data generated improves the performance of networks.

REFERENCES

Adib, A., & Mahmoodi, A. (2017). Prediction of suspended sediment load using ANN GA conjunction model with Markov chain approach at flood conditions. *KSCE Journal of Civil Engineering, 21*(1), 447-457.

Ahmed, J. A., & Sarma, A. K. (2007). Artificial neural network model for synthetic streamflow generation. *Water Resources Management, 21*(6), 1015-1029.

Aksoy, H. (2003). Markov chain-based modeling techniques for stochastic generation of daily intermittent streamflows. *Advances in Water Resources, 26*(6), 663-671.

Anmala, J., Zhang, B., & Govindaraju, R. S. (2000). Comparison of ANNs and empirical approaches for predicting watershed runoff. *Journal of Water Resources Planning and Management- ASCE, 126*(3), 156-166.

Banihabib, M. E., Ahmadian, A., & Jamali, F. S. (2017). Hybrid DARIMA-NARX model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic Oscillation (NAO) and rainfall data. *GeoResJ, 13*, 9-16.

Cigizoglu, H. K. (2003). Incorporation of ARMA models into flow forecasting by artificial neural networks. *Environmetrics, 14*(4), 417-427.

Keskin, M. E., & Taylan, E. D. (2010). Artificial intelligent models for flow prediction: a case study on Alara Stream. *Journal of Engineering Science and Design, 1*(1), 8-13.

Khan, M. Y. A., Hasan, F., Panwar, S., & Chakrapani, G. J. (2016). Neural network model for discharge and water-level prediction for Ramganga River catchment of Ganga Basin, India. *Hydrological Sciences Journal, 61*(11), 2084-2095.

Lee, D. H., & Kang, D. S. (2016). The application of the artificial neural network ensemble model for simulating streamflow. *Procidia Engineering, 154*, 1217-1224.

Ochoa-Riveria, J. C., Garcia-bartual, R., & Andreu, J. (2002). Multivariate synthetic streamflow generation using a hybrid model based on artificial neural networks. *Hydrology and Earth System Sciences, 6*(4), 641–654.

Parsaie, A., Yonesi, H., & Najafian, S. (2017). Prediction of flow discharge in compound open channels using adaptive neuro fuzzy inference system method. *Flow Measurement and Instrumentation, 54*, 288-297.

Rezaeianzadeh, M., Stein, A., & Cox, J. P. (2016). Drought forecasting using Markov chain model and artificial neural networks. *Water Resources Management, 30*(7), 2245-2259.

Raman, H., & Sunilkumar, N. (1995). Multivariate modeling of water resources time series using artificial neural networks. *Hydrological Sciences Journal, 40*(2), 145–163.

Sarlak, N., Kahya, E., & Anwar Bég, O. (2009). Critical drought analysis: case study of Göksu River (Turkey) and North Atlantic oscillation influences. *Journal of Hydrologic Engineering- ASCE, 14*(8), 795-802.

Stošić, B. D., Silva, J. R. S., Filho, M. C., & Cantalice, J. R. B. (2012). Optimizing river discharge measurements using Monte Carlo Markov Chain. *Journal of Hydrology, 450*, 199-205.

Szilagyi, J., Balint, G., & Csik, A. (2006). Hybrid, Markov Chain-based model for daily streamflow generation at multiple catchment sites. *Journal of Hydrologic Engineering- ASCE, 11*(3), 245-256.

Xu, Z. X., Schumann, A., & Brass, C. (2001). Markov autocorrelation pulse model for two sites daily streamflow. *Journal of Hydrologic Engineering- ASCE, 6*(3), 189-195.

Xu, Z. X., Schumann, A., & Li, J. (2003). Markov cross-correlation pulse model for daily streamflow generation at multiple sites. *Advances in Water Resources, 26*(3), 325-335.

Young, C. C., Liu, W. C., & Chung, C. E. (2015). Genetic algorithm and fuzzy neural networks combined with the hydrological modeling system for forecasting watershed runoff discharge. *Neural Computing and Applications, 26*(7), 1631-1643.

Pertanika J. Sci. & Technol. 26 (4): 1789 - 1800 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Reliability Analysis of Screw Manufacturing Plant Using Orthogonal Matrix Method

Shikha Bansal* and Sohan Tyagi

Department of Mathematics, SRM University, NCR Campus, Modinagar-201204, India

ABSTRACT

The paper proposes a methodology to compute reliability of the screw plant. A screw plant consists of four subsystems A, B, C, D working in series namely heading machine, slotting machine, thread rolling machine and polishing machine. Subsystem A is supported by standby units having perfect switching over device and subsystem C has two units in parallel redundancy and remaining two subsystems B and D are subjected to major failure only. For system configuration and establishment of the model we prefer Boolean algebra method and orthogonal matrix method has been used for reliability calculation. Reliability of the Screw Plant has been estimated when the failure rate expressed by weibull and exponential time distribution. Mean time to failure has also been determined for exponential time distribution, which is also a relevant characteristic of reliability.

Keywords: Boolean function, failure rate, mean time to failure, reliability

INTRODUCTION

Now a day, companies are required to place strenuous efforts . One has often heard the expression "There is no substitute for experience" and many industries today depend on this experience and skill to produce reliable products. In a progressive business surrounding, a reliable production system secures the sustainability of an enterprise. Thus, the system reliability assessment and prediction, which concerns the different stages of the operating process, has become increasingly important. In this paper, we resolve the mathematical function that explicits the system reliability in terms of the component reliabilities. Previously a lot of research work has been accomplished by many analysts on the topic reliability/availability. Agarwal (1977) discussed a method to minimize the cost and maximize the availability of the system.

ARTICLE INFO

Article history:
Received: 13 December 2017
Accepted: 05 July 2018
Published: 24 October 2018

E-mail addresses:
srbansal2008@gmail.com (Shikha Bansal)
sohanridhi@gmail.com (Sohan Tyagi)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

Gupta & Kumar and Singh (1989) had studied on availability of paper industry. Dhillon (1992) explained a K-Out of N component system with human error. Burns (1997) used a computer program to calculate the mean time between failures of complex system. Gupta et al. (2005) presented numerical analysis of reliability and availability of the Serial Processes in Butter-Oil Processing Plant. Agarwal and Bansal (2009) carried head of line repair discipline with environment effect to find the reliability of the system. Agarwal et al. (2010) studied on reliability with cold standby system. Garg et al. (2010) performed a study on the blackboard manufacturing system. Ekata and Singh (2011) developed various characteristics of reliability with Partial and Complete Failure. Shakuntala et al. (2011) had applied Lagrange's method to compute the reliability of polytube industry. Singh et al. (2011) studied on a different technique to determining terminal reliability. Kumar et al. (2013) evaluated the availability analysis of coal-fired power plant. Agarwal and Bansal (2015) examined the cost analysis of solar thermal electric power plant. Taj et al. (2017) proposed regenerative point method to find reliability of cable plant. There are such a significant number of existing strategies to acquire system reliability But all the methods follow by them bring on tiresome and complicated computation. So keeping all this fact in perspective, we used Boolean algebra technique to find a formulation of the system and applying matrix method to use reliability calculation which described the structure of complex system/equipments and the features of its. Fratta (1973) demonstrated an efficient algorithm for the analysis of unreliable communication network. Nakazawa (1977) focused on a decomposition strategy for analyzing the reliability by a Boolean Expression. Agarwal and Gupta (1983) proposed Boolean algebra method to find reliability. Singh et al. (2011) examined the utilization of the Boolean Truth Table demonstrating strategy in assessing the reliability parameters. Iqbal and Uduman (2016) established Boolean function with fuzzy logic technique for Reliability analysis of paper plant, which was obviously, an achievement in this context, but still needed further development. The aim of the present paper is a further augmentation of Boolean function technique by considering a contextual analysis of screw manufacturing plant. When we want need to get instant system's reliability without numerous computations, this technique is of much significance. Generally engineers and business personnel utilize this technique. In actuality, the significance of orthogonal matrix method for complex systems cannot be overlooked because of two reasons

(a) For system reliability various techniques are required in the advancement of complex system
(b) In several engineering systems, system configurations do not modify for a long period, however the reliabilities of parts modification perpetually.

In both these circumstances, orthogonal matrix method would incredibly decrease the calculations which are required over and over to evaluate the system reliability. In this paper, we proposed a Boolean function technique to evaluate the reliability of screw

manufacturing plant. Boolean function technique is simpler as compared with the other techniques to get reliability parameters. Some specific cases have also given to enhance practical utility of the model. We have compared the values of reliability function, in case of failures follow Weibull and Exponential time distributions

System Description

Screw is a tool that has been used since the dawn of time; it is one of the most important pieces you need to repair any aspect of a project you're looking for. The word "screw" now has a permanent place in the language that we use every day. The screw plant is divided into four subsystems, namely Cold Heading Machine, Slotting Machine, Thread Rolling Machine and Polishing Machine. Usually, screws are prepared from wire. Firstly the wire is cut in desire length and makes the head of the screw by Heading Machine. Slotted head screw needs a further Step to cut the slot in the head, this is succeeded with Slotting Machine then put all the pieces into the Thread Rolling machine to produce threads. Finally, polish the screw to shine the final product. The system configuration of the system is shown in Figure 1.

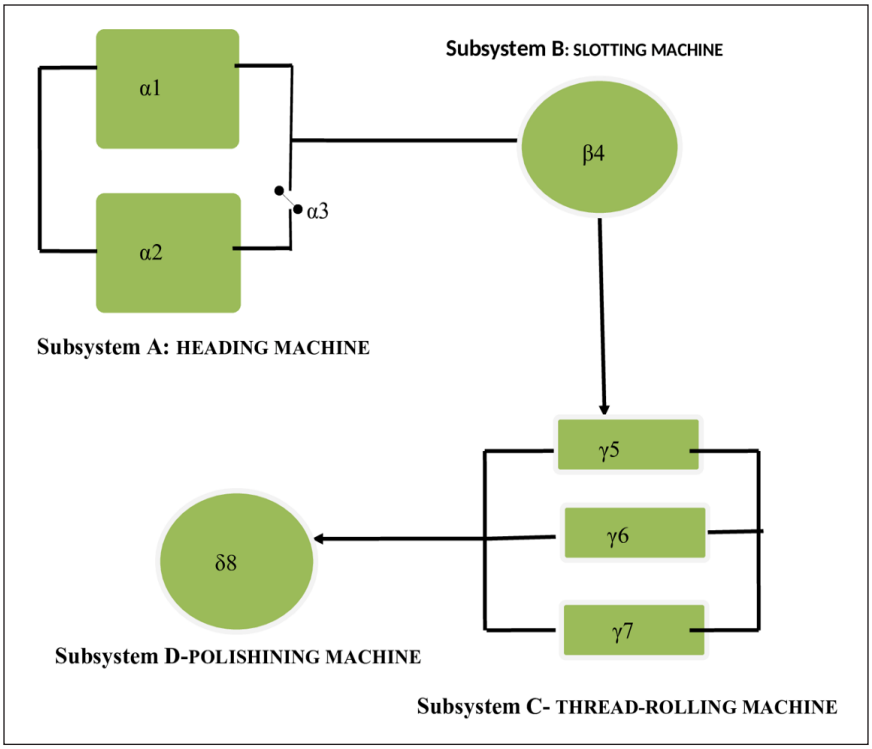


Figure 1. System configuration of screw manufacturing plant

1. Heading Machine (A) has two units; one is in operable state and second is in standby with perfect switching over device.
2. Slotting Machine (B) has a single unit.
3. Thread-Rolling Machine (C) has three units connected in parallel.
4. Polishing Machine (D) consists one unit only.

Assumptions.

1. At the beginning, all the subsystems are workable.
2. The plant has two states only i.e. Operable and failed.
3. Failure rates are arbitrary and statistically-independent.
4. The reliability of the each unit is known in advance.
5. Failure of any subsystem of the plant discontinues the functioning of the system.

Notations.

α_1, α_2	:	States of Heading machine
α_3	:	States of perfect Switching devices
β_4	:	States of Slotting machine
$\gamma_5, \gamma_6, \gamma_7$	:	States of Threading Rolling Machine
δ_8	:	States of Polishing Machine
$\alpha_{i,}\ \beta_4, \gamma_j, \delta_8$ for all i,=1,2,3 & j=5,6,7	:	=1 is in operable state and = 0 is in failed state
$\neg\alpha_i, \neg\beta_4, \neg\gamma_j\ \neg\delta_8$	:	Negation of $\alpha_{i,}\ \beta_4,\ \gamma_j, \delta_8$ for all i,=1,2,3 & j=5,6,7
$\wedge / \vee$	:	Conjunction / Disjunction
$\vert\ \vert$	:	Representation of logical matrix
$\Re_{\alpha_i}, \Re_{\beta_j}, \Re_{\gamma_k}, \Re_{\delta_m}$	:	Reliability of $\alpha_i^{\hbar}, \beta_j^{\hbar}, \gamma_k^{\hbar}, \delta_l^{\hbar}$ part of the system, $\forall$ i=1,2,3, j=4, k=5,6,7, m=8
C_i	:	$1-R_i$
$\Re_{\mathcal{SP}}$	:	Reliability of the Screw Plant
$\Re_W(t) / \Re_E(t)$	:	Reliability of the screw plant when failures follow Weibull time distribution / Exponential time distribution

Mathematical Formulation of the Model

1. There are some steps to define Boolean algebra technique
2. We discover straightforward ways between units of a plant.
3. Corresponding to the Boolean variables which are given to the diverse units, we find the Boolean expression to the paths
4. For the Boolean expression which is given in step 2, we find the disjoint expression

To get the terminal reliability, substitute the relating values of probabilities in the given disjoint Boolean expression

The Paths to refer successful activity of the screw plant by adopting the Boolean function technique, in terms of logical matrix are:

$$\Delta(\alpha_1,\alpha_2,\alpha_3,\beta_4,\gamma_5,\gamma_6,\gamma_7,\delta_8)=\begin{vmatrix} \alpha_1 & \beta_4 & \gamma_5 & \delta_8 & \\ \alpha_1 & \beta_4 & \gamma_6 & \delta_8 & \\ \alpha_1 & \beta_4 & \gamma_7 & \delta_8 & \\ \alpha_2 & \alpha_3 & \beta_4 & \delta_5 & \delta_8 \\ \alpha_2 & \alpha_3 & \beta_4 & \delta_6 & \delta_8 \\ \alpha_2 & \alpha_3 & \beta_4 & \delta_7 & \delta_8 \end{vmatrix} \quad (1)$$

Solution of the Model

With the help of algebra of logic equation (1) developed as,

$$\Delta(\alpha_1,\alpha_2,\alpha_3,\beta_4,\gamma_5,\gamma_6,\gamma_7,\delta_8)=\beta_4\,\delta_8\wedge\Psi(\alpha_1,\alpha_2,\alpha_3,\gamma_5,\gamma_6,\gamma_7) \quad (2)$$

Where,

$$\Psi(\alpha_1,\alpha_2,\alpha_3,\gamma_5,\gamma_6,\gamma_7)=\begin{vmatrix} \alpha_1 & \gamma_5 & \\ \alpha_1 & \gamma_6 & \\ \alpha_1 & \gamma_7 & \\ \alpha_2 & \alpha_3 & \gamma_5 \\ \alpha_2 & \alpha_3 & \gamma_6 \\ \alpha_2 & \alpha_3 & \gamma_7 \end{vmatrix}=|D_g| \quad (3)$$

Where, $|D_g|$ is a column matrix, $\forall$ g = 1, 2, 3,4,5,6

$$D_1=|\alpha_1 \quad \gamma_5\ | \quad (4)$$

$$D_2=|\alpha_1 \quad \gamma_6\ | \quad (5)$$

$$D_3=|\alpha_1 \quad \gamma_7\ | \quad (6)$$

$$D_4=|\alpha_2 \quad \alpha_3 \quad \gamma_5\ | \quad (7)$$

$$D_5 = \begin{vmatrix} \alpha_2 & \alpha_3 & \gamma_6 \end{vmatrix} \quad (8)$$

$$D_6 = \begin{vmatrix} \alpha_2 & \alpha_3 & \gamma_7 \end{vmatrix} \quad (9)$$

The defined logical matrix in equation (3) represents cases of parallel units. D_1 represents when α_1 and γ_5 are working. D_2 is case when α_1 and γ_6 are working and so on.

By Orthogonalization algorithm the above equation may be written as

$$\Psi(\alpha_1,\alpha_2,\alpha_3,\gamma_5,\gamma_6,\gamma_7) = \begin{vmatrix} D_1 & & & & & \\ \neg D_1 & D_2 & & & & \\ \neg D_1 & \neg D_2 & D_3 & & & \\ \neg D_1 & \neg D_2 & \neg D_3 & D_4 & & \\ \neg D_1 & \neg D_2 & \neg D_3 & \neg D_4 & D_5 & \\ \neg D_1 & \neg D_2 & \neg D_3 & \neg D_4 & \neg D_5 & D_6 \end{vmatrix} \quad (10)$$

$\neg D_1$ Represent that D_1 is in failed state. Second row $\neg D_1 D_2$ represent that D_1 is in failed state and D_2 is working. Similarly other rows can be defined. By application of Boolean algebra of logical matrix above rows can be expanded as:

$$\neg D_1\, D_2 = \begin{vmatrix} \neg\alpha_1 & \\ \alpha_1 & \neg\gamma_5 \end{vmatrix} \wedge \begin{vmatrix} \alpha_1 & \gamma_6 \end{vmatrix} = \begin{vmatrix} \alpha_1 & \neg\gamma_5 & \gamma_6 & \end{vmatrix} \quad (11)$$

Similarly,

$$\neg D_1 \neg D_2\, D_3 = \begin{vmatrix} \alpha_1 & \neg\gamma_5 & \neg\gamma_6 & \gamma_7 \end{vmatrix} \quad (12)$$

$$\neg D_1 \neg D_2 \neg D_3\, D_4 = \begin{vmatrix} \neg\alpha_1 & \alpha_2 & \alpha_3 & \gamma_5 \end{vmatrix} \quad (13)$$

$$\neg D_1 \neg D_2 \neg D_3 \neg D_4\, D_5 = \begin{vmatrix} \neg\alpha_1 & \alpha_2 & \alpha_3 & \neg\gamma_5 & \gamma_6 \end{vmatrix} \quad (14)$$

$$\neg D_1 \neg D_2 \neg D_3 \neg D_4 \neg D_5\, D_6 = \begin{vmatrix} \neg\alpha_1 & \alpha_2 & \alpha_3 & \gamma_5 & \neg\gamma_6 & \gamma_7 & \end{vmatrix} \quad (15)$$

Inserting all these values in equation (10), we get

$$\Psi(\alpha_1,\alpha_2,\alpha_3,\gamma_5,\gamma_6,\gamma_7) = \begin{vmatrix} \alpha_1 & \gamma_5 & & & & \\ \alpha_1 & \neg\gamma_5 & \gamma_6 & & & \\ \alpha_1 & \neg\gamma_5 & \neg\gamma_6 & \gamma_7 & & \\ \neg\alpha_1 & \alpha_2 & \alpha_3 & \gamma_5 & & \\ \neg\alpha_1 & \alpha_2 & \alpha_3 & \neg\gamma_5 & \gamma_6 & \\ \neg\alpha_1 & \alpha_2 & \alpha_3 & \gamma_5 & \neg\gamma_6 & \gamma_7 \end{vmatrix} \quad (16)$$

Using (16), equation (2) becomes

$$\Delta(\alpha_1,\alpha_2,\alpha_3,\beta_4,\gamma_5,\gamma_6,\gamma_7,\delta_8)=\begin{vmatrix} \alpha_1 & \beta_4 & \gamma_5 & \delta_8 & & & & \\ \alpha_1 & \beta_4 & \neg\gamma_5 & \gamma_6 & \delta_8 & & & \\ \alpha_1 & \beta_4 & \neg\gamma_5 & \neg\gamma_6 & \gamma_7 & \delta_8 & & \\ \neg\alpha_1 & \alpha_2 & \alpha_3 & \beta_4 & \gamma_5 & \delta_8 & & \\ \neg\alpha_1 & \alpha_2 & \alpha_3 & \beta_4 & \neg\gamma_5 & \gamma_6 & \delta_8 & \\ \neg\alpha_1 & \alpha_2 & \alpha_3 & \beta_4 & \gamma_5 & \neg\gamma_6 & \gamma_7 & \delta_8 \end{vmatrix} \quad (17)$$

Hence the probability of successful operation (i.e. reliability) of the screw plant is given by

$$\Re_S = P_r\{\Delta(\alpha_1,\alpha_2,\alpha_3,\beta_4,\gamma_5,\gamma_6,\gamma_7,\delta_8)=1\}$$

Or,

$$\Re_S = \Re_{\alpha_1}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\delta_8} + \Re_{\alpha_1}\Re_{\beta_4}C_{\gamma_5}\Re_{\gamma_6}\Re_{\delta_8} + \Re_{\alpha_1}\Re_{\beta_4}C_{\gamma_5}C_{\gamma_6}\Re_{\gamma_7}\Re_{\delta_8} + C_{\alpha_1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}C_{\delta_8}$$
$$+ C_{\alpha_1 1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}C_{\gamma_5}\Re_{\gamma_6}\Re_{\delta_8} + C_{\alpha_1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}C_{\gamma_6}\Re_{\gamma_7}\Re_{\delta_8}$$

$$\Re_S = \Re_{\delta_8}\Big[\Re_{\alpha_1}\Re_{\beta_4}\Re_{\gamma_5} + \Re_{\alpha_1}\Re_{\beta_4}\Re_{\gamma_6} - \Re_{\alpha_1}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_6} + \Re_{\alpha_1}\Re_{\beta_4}\Re_{\gamma_7} - \Re_{\alpha_1}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_7}$$
$$- \Re_{\alpha_1}\Re_{\gamma_5}\Re_{\gamma_6}\Re_{\gamma_7} + \Re_{\alpha_1}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_6}\Re_{\gamma_7} + \Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5} - \Re_{\alpha_1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5} + \Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_6}$$
$$- \Re_{\alpha_1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_6} - \Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_6} + \Re_{\alpha_1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_6} + \Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_7}$$
$$- \Re_{\alpha_1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_7} - \Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_6}\Re_{\gamma_7} + \Re_{\alpha_1}\Re_{\alpha_2}\Re_{\alpha_3}\Re_{\beta_4}\Re_{\gamma_5}\Re_{\gamma_6}\Re_{\gamma_7}\Big]$$

(18)

Some Special Cases

Case I: If $\Re_{\alpha_i} = \Re_{\beta_4} = \Re_{\gamma_j} = \Re_{\delta_8} \quad \forall i = 1,2,3 \quad and \quad j = 5,6,7$

Then equation (18) turns into

$$\Re_S = \Re^8 - \Re^7 - \Re^6 - \Re^5 + 3\Re^4 \quad (19)$$

Case II: In case of weibull time distribution let h_i be the failure rate corresponding to system state $\alpha_i, \forall i = 1,2,3,\ \beta_4,\ \gamma_j\ \ \forall j = 5,6,7,\ \delta_8$ respectively, then at an instant 't' the system reliability is given by

$$\Re_W(t) = \sum_{k=1}^{9} \exp.\{-P_k t^{\theta}\} - \sum_{l=1}^{8} \exp.\{-N_l t^{\theta}\} \quad (20)$$

Where,

$$P_1 = h_{\alpha_1} + h_{\beta_4} + h_{\delta_8} + h_{\delta_8}$$

$$P_2 = h_{\alpha_1} + h_{\beta_4} + h_{\gamma_6} + h_{\delta_8}$$

$$P_3 = h_{\alpha_1} + h_{\beta_4} + h_{\gamma_7} + P_4$$

$$P_4 = h_{\alpha_1} + h_{\alpha_2} + h_{\beta_4} + h_{\gamma_5} + h_{\delta_8}$$

$$P_5 = h_{\alpha_1} + h_{\alpha_2} + h_{\beta_4} + h_{\delta_8} + h_{\delta_8}$$

$$P_6 = h_{\alpha_1} + h_{\beta_4} + h_{\gamma_5} + h_{\gamma_6} + h_{\gamma_7} + h_{\delta_8}$$

$$P_7 = h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_5} + h_{\gamma_7} + h_{\delta_8}$$

$$P_8 = h_{\alpha_1} + h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_6} + h_{\gamma_6} + h_{\delta_8}$$

$$P_9 = h_{\alpha_1} + h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_5} + h_{\gamma_6} + h_{\gamma_7} + h_{\delta_8}$$

and

$$N_1 = h_{\alpha_1} + h_{\beta_4} + h_{\gamma_5} + h_{\gamma_6} + h_{\delta_8}$$

$$N_2 = h_{\alpha_1} + h_{\beta_4} + h_{\gamma_5} + h_{\gamma_7} + h_{\delta_8}$$

$$N_3 = h_{\alpha_1} + h_{\beta_4} + h_{\gamma_6} + h_{\gamma_7} + h_{\delta_8}$$

$$N_4 = h_{\alpha_1} + h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_5} + h_{\delta_8}$$

$$N_5 = h_{\alpha_1} + h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_6} + h_{\delta_8}$$

$$N_6 = h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_5} + h_{\gamma_6} + h_{\delta_8}$$

$$N_7 = h_{\alpha_1} + h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_7} + h_{\gamma_7} + h_{\delta_8}$$

$$N_8 = h_{\alpha_2} + h_{\alpha_3} + h_{\beta_4} + h_{\gamma_5} + h_{\gamma_6} + h_{\gamma_7} + h_{\delta_8}$$

Case III: Consider $\theta = 1$, in the case of exponential time distribution. It is a special case of the weibull distribution. At an instant‘t’ the reliability of considered system is given by:

$$\Re_E(t) = \sum_{k=1}^{9} \exp.\{-P_k t\} - \sum_{l=1}^{8} \exp.\{-N_l t\} \quad (21)$$

Where, P_k and N_l 's have previously described.

Also mean time to failure of the screw plant in case of exponential time distribution is given by

$$\text{M.T.T.F} = \int_0^\infty \Re_E(t)\, dt$$

$$= \sum_{k=1}^{9}\left(\frac{1}{P_k}\right) - \sum_{l=1}^{8}\left(\frac{1}{N_l}\right) \qquad (22)$$

Numerical Calculation

For justification of above calculate expression we have to set some appropriate numerical data and result will be given below. The relationship can be seen by tabular and graphical representation.

(A) $\alpha_i(i=1,2,3), \beta_4, \gamma_j(j=5,6,7), \delta_8 = 0.001, \quad \theta = 2 \; and \;$ t = 0,1,2 - - in equation (20)

(B) $\alpha_i(i=1,2,3), \beta_4, \gamma_j(j=5,6,7), \delta_8 = 0.001, \quad and \;$ t = 0,1,2 - - in equation (21)

(C) $\alpha_i(i=1,2,3), \beta_4, \gamma_j(j=5,6,7), \delta_8 = 0,0.1.....1.0$ in equation (22)

With the help of above data, we figure out the Table 1 and Table 2 .The analogous graph has been shown in Figure 2 and Figure 3 respectively.

RESULT AND DISCUSSION

In the compiled work of this paper, the reliability of the screw manufacturing plant has been carried out. The reliability which depict the performance of the plant, when failure rates follow exponential and Weibull time distribution with respect to the time is shown by tabular and graphical representation in the Tables 1 and 2 and Figures 2 and 3 respectively. Figure 1 i.e. 'Reliability Vs Time" highlight that when failure rate follows an exponential time distribution, the reliability of the screw plant decreases approximately at a uniform rate, but in case of weibull time distribution it decreases speedily. The study of Table 2 and Figure 2 "MTTF V/S Failure Rate" explores that the mean time to failure of the plant decreases catastrophically in the starting but later it decreases approximately at a uniform rate.

Table 1
Comparison of reliability when failures follow Weibull and Exponential time distribution

Time t (in days)	RW(t)	RE(t)
0	1	1
1	0.998001	0.998001
2	0.992016	0.996004
3	0.982079	0.994009
4	0.968244	0.992016
5	0.950581	0.990025
6	0.929175	0.988035
7	0.904122	0.986048
8	0.875537	0.984062
9	0.843558	0.982079
10	0.808362	0.980097

Table 2
Mean Time to Failure (MTTF) corresponding to failure rate

Failure, f_i	MTTF
0.01	36.54762
0.02	18.27381
0.03	12.18254
0.04	9.136905
0.05	7.309524
0.06	6.09127
0.07	5.221088
0.08	4.568452
0.09	4.060847
0.1	3.654762

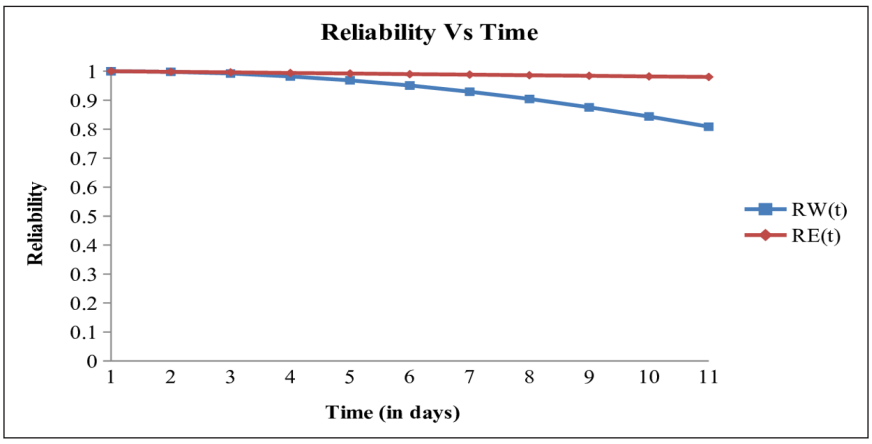


Figure 2. MTTF vs. failure rate

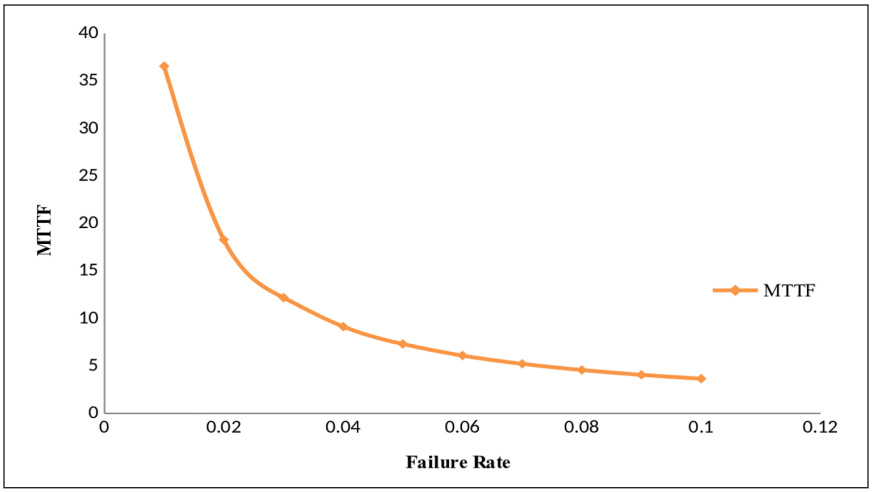


Figure 2. MTTF vs. failure rate

REFERENCES

Agarwal, K. K. (1977). Minimum cost system with specified reliability. *IEEE Transactions on Reliability, 26*(3), 166-167.

Agarwal, S. C., & Bansal, S. (2009). Reliability analysis of a standby redundant complex system with changing environment under head of line repair discipline. *Bulletin of Pure and Applied Sciences, 28*(1), 165-173.

Agarwal, S. C., & Bansal, S. (2015). Cost analysis of solar thermal electric power plant. *International Journal of Advanced Technology in Engineering and Science, 3*(10), 12-22.

Agarwal, S. C., Sahani, M., & Bansal, S. (2010). Reliability characteristic of cold-standby redundant system. *International Journal of Research and Review in Applied Sciences (IJRRAS), 3*(2), 193-199.

Burns, R. J. (1997). A simple technique for obtaining the MTBF of complex systems through numerical integration of the reliability function. *Microelectronics Reliability, 37*(3), 397-401.

Dhillon, B. S. (1992). Stochastic modeling of k-out of n units family of system. *International Journal of System Science, 23*(8), 1277-1287.

Dragovic, I., Turajlic, N., Radojevic, D., & Petrovic, B. (2014). Combining Boolean consistent fuzzy logic and ahp illustrated on the web service selection problem. *International Journal of Computational Intelligence Systems, 7*(1), 84-93.

Ekata, S. S. (2011). Operational readiness of global mobile satellite communication system under partial and complete failure. *Journal of Engineering Science and Technology, 6*(1), 82-93.

Fratta , F., & Montanari , U. G. (1973). A Boolean algebra method for computing the terminal reliability in a communication network. *IEEE Transactions on Circuit Theory, 20*(3), 203-211.

Garg, S., Singh, J., & Singh, D. V. (2010). Availability and maintenance scheduling of a repairable blackboard manufacturing system. *International Journal of Reliability and Safety*, *4*(1), 497-508.

Gupta, P. P., & Agarwal, S. C. (1983). A Boolean algebra method for reliability calculations *Microelectronics Reliability, 23*(5), 863-865.

Gupta, P., Lal, A. K., Sharma, R., & Singh, J. (2005). Numerical analysis of reliability and availability of the serial processes in butter-oil processing plant. *International Journal of Quality & Reliability Management, 22*(3), 303–316.

Iqbal, P., & Uduman Sehik, P. S. (2016). Reliability analysis of paper plant using Boolean function with fuzzy logic technique. *International Journal of Applied Engineering Research*, *11*(1), 573-577.

Kumar, D., & Singh, J. (1989). Availability of washing system in paper industry. *Microelectronics Reliability, 29*(5), 775-778.

Kumar, R., Sharma, A., & Tewari , P. C . (2013). Performance evaluation of a coal -fired power plant. *International Journal of Performability Engineering, 9*(4), 455-461.

Nakazawa, H. (1977). A decomposition method for computing system reliability by a Boolean expression. *IEEE Transactions on Reliability, 26*(4), 250-252.

Shakuntala, S., Lal, A. K., Bhatia, S. S., & Singh, J. (2011). Reliability analysis of polytube industry using supplementary variable technique. *Applied Mathematics and Computation, 218*(8), 3981-3992.

Singh, H., Vaithilingam, S., Anne, R. K., & Anneberg, L. (1996). Terminal reliability binary decision diagrams. *Microelectronics Reliability, 36*(3), 363-365.

Singh, P., Qamar, S., & Yadav, K. P. (2011). Reliability forecast for UPS (Uninterruptible Power Supply) by employing Boolean truth table method. *Journal of Pure and Applied Science & Technology, 1*(2), 66-76.

Taj, S. Z., Rizwan, S. M., Alkali, B. M., Harrison, D. K., & Taneja, G. L. (2017). Reliability analysis of a single machine subsystem of a cable plant with six maintenance categories. *International Journal of Applied Engineering Research, 12*(8), 1752-1757.

Pertanika J. Sci. & Technol. 26 (4): 1801 - 1822 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Named Entity Recognition in Hindi Using Hyperspace Analogue to Language and Conditional Random Field

Arti Jain* and Anuja Arora

CSE & IT, Jaypee Institute of Information Technology, Noida, U.P. 201309, India

ABSTRACT

Named Entity Recognition (NER) is defined as identification and classification of Named Entities (NEs) into set of well-defined categories. Many rule-based, machine learning based, and hybrid approaches have been devised to deal with NER, particularly, for the English language. However, in case of Hindi language several perplexing challenges occur that are detailed in this research paper. A new approach is proposed to perform Hindi NE Recognition using semantic properties to handle some of the Hindi language specific NER challenges. And because of increasing demand in Hindi health care applications, Hindi Health Data (HHD) is crawled from four well-known Indian websites: Traditional Knowledge Digital Library; Ministry of Ayush; University of Patanjali; and Linguistic Data Consortium for Indian Languages. Four novel NE types are determined, namely- Person NE, Disease NE, Symptom NE and Consumable NE. For training purpose, HHD data is converted into Hyperspace Analogue to Language (HAL) vectors, thereby, maps each word into a high dimensional space. Conditional Random Field model is applied based on HHD feature engineering, HHD gazetteers and HAL. Blind test data is then mapped into the high dimensional space created during the training phase and outputs the annotated test data. The results obtained are quite significant; and HAL accompanied with CRF approach seems to provide effective outcome for Hindi NE Recognition.

Keywords: Conditional Random Field, Hindi, Hyperspace Analogue to Language, Named Entity Recognition

INTRODUCTION

Named Entity Recognition (NER) (NER) (Nadeau & Sekine, 2007; Ekbal & Bandyopadhyay, 2008; Srivastava et al., 2011; Rodriquez et al., 2012; Marrero et al., 2013; Baldwin et al., 2015; Ekbal et al., 2016; Patil et al., 2016; Baksa et al., 2017) is a non-trivial, automated sequence

ARTICLE INFO

Article history:
Received: 09 January 2018
Accepted: 05 May 2018
Published: 24 October 2018

E-mail addresses:
ajain.jiit@gmail.com (Arti Jain)
'anuja.arora29@gmail.com (Anuja Arora)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

labelling task which comprises identification and classification of Named Entities (NEs). Identification of NE means marking the presence of a word/term/phrase i.e. names (noun/noun phrase) as NE in a given text. And, classification of NE means denoting the role of an identified NE into certain well-defined categories such as Person, Location, Organization, Money, Date and Time. NER is treated as a main sub-task of Information Extraction (IE) (Grishman, 1995; Chinchor & Robinson, 1997) and is successful in vivid application areas such as Question Answering (Khalid et al., 2008), Machine Translation (Aggarwal & Zhai, 2012), Automatic Text Summarization (Gupta & Lehal, 2011), Word Sense Disambiguation (Moro et al., 2014) and so on. In general, there are three main approaches to NER systems, namely- Rule-based approach (Farmakiotou et al., 2000; Chiticariu et al., 2010), Machine learning approach (Jiang et al., 2011; Ekbal et al., 2016), and Hybrid approach (Saha et al., 2008; Rocktäschel et al., 2012). Rule based NER approach comprises of language based hand-crafted rules and other heuristics e.g. set of patterns to classify words for NER system. For this purpose, thorough language knowledge, grammatical expertise and advanced skills related to the language are required to achieve good results. But these rules are non-transferable to other languages and domains. Also, they incur steep maintenance cost especially when new rules are introduced for certain new information or new domain. Machine learning (ML) based NER approach requires huge amount of NE annotated training data to acquire good results. ML further comprises three approaches- Supervised learning (SL), Semi-supervised learning (SSL) and Unsupervised learning (UL). SL involves learning to classify a given set of labelled examples that are made up of the number of features, only when large amount of high quality training data is available e.g. Hiddden Markov Model (Zhou & Su, 2002), Maximum Entropy (Curran & Clark, 2003), Conditional Random Field (Ekbal & Bandyopadhyay, 2009), Support Vector Machine (Saha et al., 2010), and Decision Tree (Szarvas et al., 2006). SSL involves technique such as bootstrapping (Kozareva, 2006) which has a small degree of supervision for starting the learning process. UL involves training with few seed lists and large unannotated corpus where NEs are gathered from cluster groups based on the similarity of context. For example, Collins and Singer (1999) had discussed an unsupervised model for NE classification by the use of unlabelled examples of data. Kim et al. (2002) had proposed an unsupervised NE classification and ensemble technique which used small scale NE dictionary and unlabelled corpus for NEs.

So far, NER system for English language (Grishman, 1995; Rodriquez et al., 2012; Marrero et al., 2013) has already been widely explored. Konkol et al. (2015) had discussed latent semantic based information for NER which considered local context methods and global context methods. Local context uses only a limited context (context window) around the word to infer vector. Most prominent local context methods are: Hyperspace Analogue to Language (HAL), Correlated Occurrence Analogue to Lexical Semantic (COALS),

Random Indexing (RI), Bound Encoding of AggreGate Language Environment (BEAGLE), Purandare and Pedersen (P&P). While global context uses a wider context (whole section or document) around the word to infer vector. Widely used global context methods are: Latent Semantic Analysis (LSA), Probabilistic Latent Semantic Analysis (PLSA), Latent Dirichlet Allocation (LDA). But these latent semantic based methodologies have been applied for NER in English and some European languages, which can further be explored for Hindi- an Indian language based NER as well. Since NER system for Hindi (Ekbal & Bandyopadhyay, 2009; Saha et al., 2010; Srivastava et al. 2011; Athavale et al., 2016) is still quite challenging.

Formulation of Problem for Hindi NER

Hindi is written in the Devanagari (Gupta et al., 2011) script and is considered as an official language of the Government of India, in addition to English. And outside India, it is an official language in Fiji, and regional language in Mauritius, Trinidad and Tobago, Guyana, and Suriname. Hindi is highly inflectional, morphologically rich and primarily suffixing language. An excellent source for Hindi language processing is the Hindi WordNet (http://www.cfilt.iitb.ac.in/wordnet/webhwn/). From the past few years, Hindi NER task (Cucerzan & Yarowsky, 1999; Li & McCallum, 2003; Ekbal et al., 2008; Saha et al., 2008; Krishnarao et al., 2009; Srivastava et al., 2011; Athavale et al., 2016) is considered as a budding research topic.

In this paper, NER for the Hindi language using machine learning based Hyperspace Analogue to Language (HAL) methodology is proposed. System is trained for Hindi Health Domain (HHD) corpus, large part of which is unlabelled, as input, which is then transformed into feature vectors (representation of words) along with labels (representation of entities). To do so, training data passes through pre-processing (tokenization, fill-in missing values or gaps); HAL steps (word-vector generation, co-occurrence matrix, and similarity measurement); feature engineering (head nouns, word suffix, part-of-speech and n-gram); seed gazetteers and their extensions through Hindi WordNet. Training algorithm then estimates parameters for the Conditional Random Filed (CRF) model using the trained data. After completion of the learning process, the unannotated blind HHD test data is processed and is transformed into feature vectors using HAL model. CRF is then applied and map the test words to HHD NEs. System implementation of the proposed Hindi NER approach is performed using python® (https://www.python.org/downloads/release/python-2711/) and its supportive ML libraries, to develop HAL model and to find NEs for the test data using CRF. The results that are obtained are quite significant, and the proposed approach is novel for the Hindi NER system.

While performing the Hindi language based NER task, several implementation challenges are encountered. The biggest challenge was crawling of Hindi health corpus

from various websites which was quite complex as Hindi language-based sites followed varying Unicode styles and required integration. Lack of certain python libraries to Unicode support; compatibility of Hindi WordNet w.r.to.python; formulation of gazetteer lists due to lack of standardized Hindi health-based gazetteers are few other challenges. The proposed NER approach is applicable on any social media such as Twitter health tweets, Patanjali Ayurvedic site etc., wherever health content is made available in Hindi.

Research Contributions

This research provides three valuable contributions as stated below:

RC1: Explore Hindi health domain based named entities and relevant gazetteers using Hindi lexical resource (Hindi WordNet).

RC2: Propose latent semantics-based Hyperspace Analogue to Language as state-of-art NER technique for Hindi.

RC3: Study the impact of feature engineering, gazetteers and HAL on NER in Hindi and achieve significant results using CRF method.

Organization of Paper

The rest of the paper is organized as follows: Section 2 introduces the chosen training corpus from Hindi health domain (HHD) Indian sites. This section comprises of HHD NEs and HHD gazetteers. Section 3 discusses contemporary challenges in Hindi NER. Section 4 gives detailed architecture of proposed NER system for Hindi. Section 5 illustrates hyperspace analogue to language semantic details using HAL algorithm and HAL example. Section 6 explains machine learning based CRF model. Section 7 describes HHD feature engineering module. Section 8 shows experiments and results of the Hindi NE Recognition system. Finally, Section 9 concludes the paper.

Training Corpus

Due to the growing need of smart health applications (www.onlymyhealth.com; https://pmsma.nhp.gov.in/) in Hindi there is a rapid demand for health related NER system. As far, no standard Hindi health-domain corpus is available, so we have crawled data of 310,530 words from the four well-known Hindi health domain based Indian websites viz. (i) Traditional Knowledge Digital Library (http:// www.tkdl.res.in/), (ii) Ministry of Ayush (http://ayush.gov.in/), (iii) University of Patanjali (https://www.patanjaliayurved.net/), and (iv) Linguistic Data Consortium for Indian Languages (http://www.ldcil.org/). Figure 1 represents sample crawled Hindi Health Domain (HHD) corpus.

In this research work, we have considered four NEs- Person (PER), Disease (DIS), Symptom (SMP) and Consumable (CNS) NEs for HHD corpus. As per the NER research

किसी एक बीमारी से हमारे देश में इतनी मौतें नहीं होतीं जितनी निमोनिया (Pneumonia) से होती हैं।

निमोनिया से बचाव और इसका इलाज बेहद सुगम है लेकिन अकसर लोगों के पास इसके जुड़ी जानकारी नहीं होती।
जानिए निमोनिया के विषय में सभी बातें।

निमोनिया के घरेलू उपचार

- हल्दी, काली मिर्च, मेथी और अदरक जैसे प्रतिदिन उपयोग में आने वाले खाद्य प्रदार्थ फेफड़ों के लिए फायदेमंद होते हैं।
- तिल के बीज भी निमोनिया के उपचार में सहायक होते हैं। 300 मिलीलीटर पानी में 15 ग्राम तिल के बीज, एक चुटकी साधारण नमक, एक चम्मच

- अलसी और एक चम्मच शहद मिलकर प्रतिदिन उपयोग करने से फेफड़ों से कफ बाहर निकलता है।

- ताजा अदरक का रस लेने या अदरक को चूसने से भी निमोनिया में आराम मिलता है।

- थोड़े से गुनगुने पानी के साथ शहद लेना भी लाभदायक रहता है।

- गर्म तारपीन तेल का और कपूर के मिश्रण से छाती पर मालिश करने से निमोनिया से राहत मिलती है।

- रोगी का कमरा स्वच्छ, और गर्म होना चाहिए। कमरे में सूर्य की रौशनी अवश्य आनी चाहिये।

- रोगी के शरीर को गर्म रखें, विशेषकर छाती और पैरों को।

Figure 1. HHD sample crawled data

guidelines that are undertaken by the AU-KBC Research Centre, Chennai, among the four considered NE types, the first two NEs-PER and DIS are direct sub-categories of ENAMEX (ltrc.iiit.ac.in/iasnlp2014/slides/lecture/sobha-ner.ppt), the third NE- SMP is extracted from fine-grained variation of DIS, and the fourth NE- CNS is extracted from Material sub-category of ENAMEX. Further, the current work can be extended with some more NEs that can be made available in the chosen corpus such as Food, Diagnosis, Treatment etc. Presently, all such words are considered as Not-Named Entity (NNE). Figure 2 shows the considered NEs and their integration relationships. Although researchers in the past have identified Person NE in news-wire domain but they have not worked with the rest three NEs because there is no research work being conducted so far on health domain for Hindi NER.

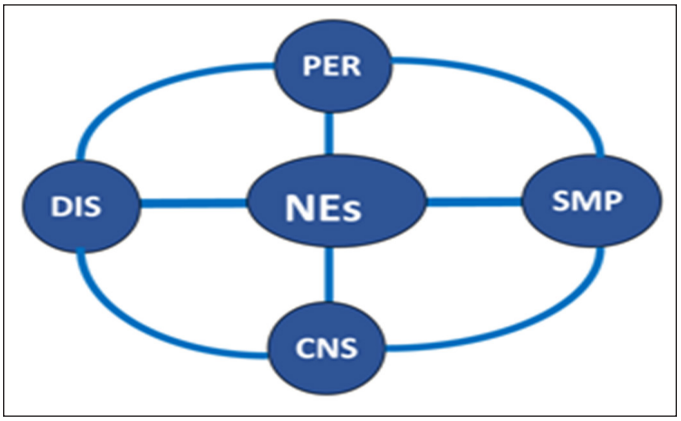


Figure 2. HHD named entities

Description of the selected four NE types for this NER research is given as follows:

Person (PER): PER refers to the person (human being) who may be a single individual or a group. Person entity in HHD corpus contains semantic roles of a person who is directly or indirectly involved in or is affected by certain disease. For Example- "व्यक्ति" (Vyakti/Person), "मरीज़" (Marij/Patient), "महिला" (Mahila/Lady), "आदमी" (Aadmi/Man), "चिकित्सक" (Chikitsak/Doctor), "एक्सपर्ट्स" (Experts) etc.

Disease (DIS): DIS refers to the name of the disease that adversely affects a patient irrespective of the fact that the disease is mild or severe For Example- "दमा" (Dama/Asthama), "चेचक" (Chechak/Chickenpox), "हैजा" (Heja/Cholera), "कालीखांसी" (Kali Khasi/Whooping Cough) etc.

Symptom (SMP): SMP refers to an undesirable physical or mental state of a patient that is regarded as an indicator to some well-known or unknown disease. For Example- "सूजन" (Sujan/Swelling), "मतली" (Matli/Nausea), "पीड़ा" (Peeda/Pain), "इंफेक्शन" (Infection), "तकलीफ" (Takleef/Problem), "थकावट" (Thakawat/Tiredness) etc.

Consumable (CNS): CNS refers to a substance that person intakes through various modes (e.g. oral, inject, inhale, drink, suck, swallow, eat, chew etc.) and is used in pharmacology for diagnosis, prevention, cure or treatment of diseases. For Example- "लहसुन" (Lahasun/Garlic), "दूध" (Dudh/Milk), "एंटीबॉयटिक" (Antibiotic), "ग्लूकोज" (Glucose), "अनाज" (Anaj/Cereals), "राजमा" (Rajama/Beans) etc.

HHD Gazetteers

Gazetteers or gazetteer lists are the entities dictionaries which are important for performing NER effectively (Kazama & Torisawa, 2008; Dey & Prukayastha, 2013; Sahin et al., 2017). They are neither dependent on previously discovered tokens nor on annotations. They only expect a raw text as an input and then find matches based on its contents. In the current work, initial four seed gazetteer lists are chosen manually from HHD corpus. These four lists are having- 107 entries for Person NE e.g. "बच्चा" (Baccha/Child), 141 entries for Disease NE e.g. "गठिया" (Gathiya/Arthiritis), 223 entries for Symptom NE e.g. "दर्द" (Dard/Pain), and 388 entries for Consumable NE (CNS) e.g. "खाना" (Khana/Food) respectively.

Later on, each of these lists is extended through semi-automatic process using Hindi WordNet synset through python®code. As a result, each of these four gazetteers are extended, having 860 entries for PER (e.g. "बच्चा" (Baccha/Child) has extensions as "नवजात_शिशु" (Navajat_shishu/Newborn baby), "नवजातक" (Navajataka/New born), "लड़का"

(Ladka/Boy), "बालक" (Balak/Boy), "छोकड़ा" (Chhokadaa/Man-child), "छोरा" (Chhora/Lad), "छोकरा" (Chhokara/Chap), "लौंडा" (Launda/Bugger), "वत्स" (Vats/Child), "नन्हा-मुन्ना" (Nanha-munna/Child), "नन्हा_मुन्ना" (Nanha_munna/Child), "पुत्र" (Putr/Son), "बेटा" (Beta/Son), "सुत" (Sut/Son), "शिशु" (Shishu/Baby) etc.); 597 entries for DIS (e.g. "गठिया" (Gathiya/Arthritis) has extensions as "संधिवात" (Sandhivata), "संधिशोथ" (Sandhishoth), "सन्धिवात" (Sandivata), "सन्धिशोथ" (Sandhishoth), "संधि_शूल" (Sandhi_Shula), "डमरुआ" (Damruaa), "डबरुआ" (Dabruaa), "पवन-व्याधि" (Pawan-vyadhi), "आर्थ्राइटिस" (Arthritis), "आर्थराइटिस" (Arthritis) etc.); 2655 entries for SMP (e.g. "दर्द" (Dard/Pain) has extensions as "तकलीफ" (Takleef), "दरद" (Darad), "पीड़ा" (Peeda), "तक़लीफ़" (Takleef), "पीर" (Pir), "हूक" (Huuk), "उपताप" (Uptap), "उत्ताप" (Utap), "पीरा" (Pira), "वेदना" (Vedana), "बेदना" (Bedana), "क्लेश" (klesh), "व्यथा" (Vyatha), "अनुसाल" (Anusal) etc.); and 3828 entries for CNS (e.g. "खाना" (Khana/Food) has extensions as "खाद्य_वस्तु" (Khady_vastu), "खाद्य_पदार्थ" (Khady_padarth), "आहार" (Ahar), "खाद्य" (Khady), "भोज्य_पदार्थ" (Bhojy_padarth), "खाद्य_सामग्री" (Khady_samagri), "अन्न" (Ann), "आहर" (Aahr), "फूड" (Food), "भोजन" (Bhojan), "रसोई" (Rasoi), "रोटी" (Roti), "डाइट" (Diet) etc.) respectively.

Contemporary Challenges in HHD NER

NER for humans appear to be straightforward as most of the NEs are the proper names. But for a machine to learn and understand NER is comparatively hard, especially for Hindi. A few researchers have identified challenges in Hindi NER (Ekbal et al., 2016; Jain et al., 2014; Saha et al., 2012; Srivastava et al., 2011). Some new and previously mentioned Hindi NER challenges are listed below:

Rare occurrence of certain NEs in HHD corpus: e.g. "कुटकी चिरौता" (Kutki Chirauta) which is a NE under CNS and has rare occurrence in HHD corpus.

Multiple ways of mentioning the same NE:

Variation in PER semantic information: e.g. "रोगी" (Rogi), "पेशंट" (Patient), "पेशेंट" (Patient), "मरीज" (Marij), "मरीज़" (Marij) all refer to PER (Patient).

Variation in DIS semantic information: e.g. "डायबीटीज़" (Diabetes), "डायबीटीज" (Diabetes), "डायबिटीज़" (Diabetes), "डायबिटीज" (Diabetes), "मधुमेह" (Madhumeh), "मधुप्रमेह" (Madhuprameh), "इक्षुप्रमेह" (Ikshuprameh), "मूत्रकृच्छ" (Mutrakarachchh) all refer to the DIS (Diabetes).

Variation in SMP semantic information: e.g. "झुनझुनी" (Jhunjhuni), "झनझनाहट"

(Jhnajhannahat), "झुनझुनाहट" (Jhunjhunahat), "सुरसुरी" (Sursuri), "सनसनाहट" (Sansanahat), "सनसनी" (Sansani), "सनसन" (Sansan), "सन-सन" (San-san) all refer to SMP (Tingle).

Variation in CNS semantic information: e.g. "कॉफ़ी" (Coffee), "कॉफी" (Coffee), "काफी" (Coffee), "काफ़ी" (Coffee), "कॉफ़ीपावडर" (Coffee Powder), "कॉफीपावडर" (Coffee Powder), "काफीपावडर" (Coffee Powder), "काफ़ीपावडर" (Coffee Powder) all refer to CNS (Coffee).

Disease vs. Symptom: e.g. "बदहजमी" (Badhazmi/Indigestion), "ज़ुकाम" (Jukam/Colds) refer to DIS NE or SMP NE.

Lack of Capitalization: English language uses capitalization as a discriminating feature for classifying words as NEs. On the other hand, Hindi does not have the concept of capitalization at all. For example, Tuberculosis (T.B.) is a Disease in English and is represented as "टी. बी." (T.B.) in Hindi. Similarly, "एड्स" (AIDS), "विटामिनई" (Vitamin E) etc.

Lack of well-defined Gazetteers: Well-defined NE gazetteers are not freely available for Hindi.

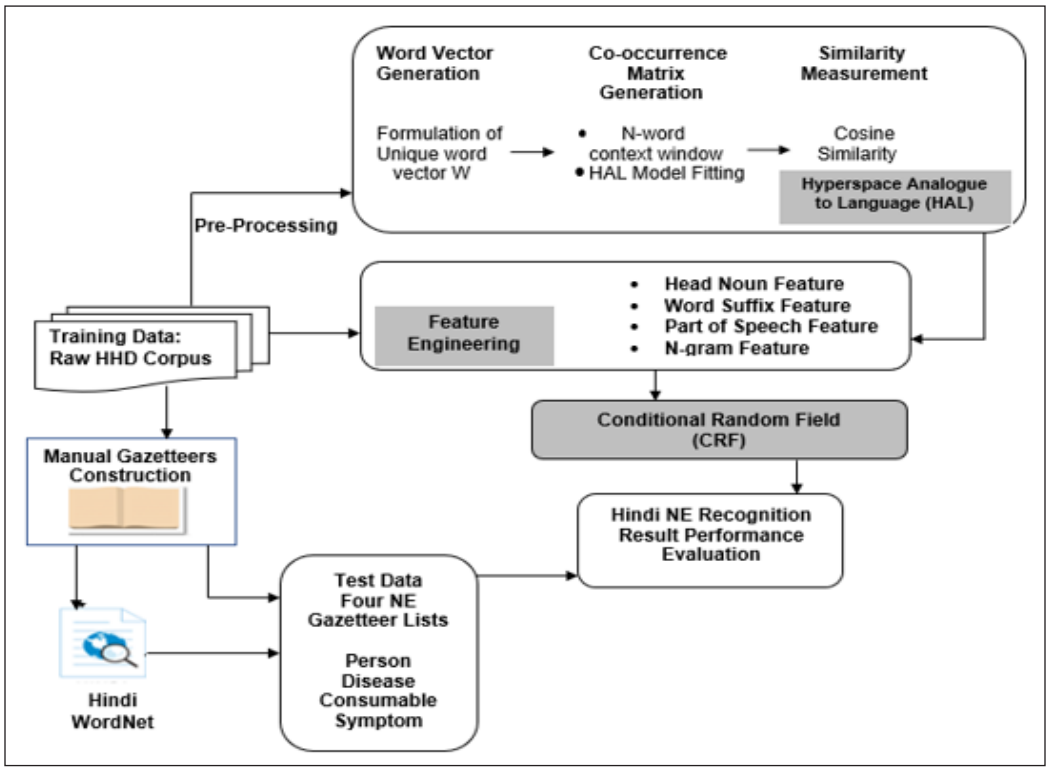


Figure 3. NER system architecture

Proposed NER System for Hindi

Architecture of HHD based NER system is depicted in Figure 3. This system works into training and test phases as follows. The training phase takes the annotated training HHD corpus and is then transformed into feature vectors (representation of words) along with labels (representation of entities). For this purpose, training data passes through HAL steps (word vector generation, co-occurrence matrix, and similarity measurement); feature engineering (head nouns, word suffix, part-of-speech and n-gram); manual gazetteer construction and their extension through Hindi WordNet. The training algorithm then estimates parameters for the CRF model using this trained dataset. Now the unannotated test data is supplied to the NER system and is transformed into feature vectors, CRF is applied onto this data and results into output annotation for the test data.

We have used nltk-3.2.4 (https://pypi.python.org/pypi/nltk), hal-0.0.3 (https://pypi.python.org/pypi/HAL), scikit-learn 0.15.2 (https://pypi.python.org/pypi/scikit-learn/0.15.2) as set of python modules for the Hindi NER task.

Hyperspace Analogue to Language

Hyperspace Analogue to Language (HAL) (Tayal et al., 2015) is also known as semantic memory and was developed by Kevin Lund and Curt Burgess, University of California, Riverside, California in 1996. HAL basic premise (Lund & Burgess, 1996; Lund et al., 1996; Burgess & Lund, 1997) relies on the fact that the words with similar meanings repeatedly occur closely (also known as co-occurrence). Another researcher (Firth, 1957) stated that a word is characterized by the company it keeps i.e. meaning of a word can be interpreted by its surrounding contexts, around which that word often appears. In this research, HAL is chosen as computational model that exploits statistics for the contexts of HHD corpus words.

HAL determines the similarities between the words while collecting the statistics about the word co-occurrences, using two vital assumptions-

Left and right context of a word holds different information, and so it is important to keep this statistic as separate entity;

Distance between the words within a sentence is important, and so more distant words are less informative while lesser distant words are more informative.

Such statistics is useful to generate high-dimensional vectors, where each vector represents meaning for one word; and the words that are represented as vectors formulate the vector space model. Then compare the words and their meanings using similarity/distance among vectors. For this purpose, HAL uses local context, also called as limited context or context window around a word to infer its vector. Such a context window contains only a few words before and after the processed word. Thus, HAL is treated as a semantic space model which discovers different kind of relations between words. Consider

an example word "रोग"(Rog/Disease) then HAL aids in finding the local context for this word as- "व्याधि" (Vyadhi/Illness) and "बीमारी"(Bimari/Disease) words that are observed to be the most similar words corresponding to the given example word.

HAL Algorithm

This section discusses about HAL algorithm and its execution through an illustrative example.

Algorithm: HAL Algorithm //python nltk, sklearn implementation

Input Hindi HHD Corpus
Output Latent semantics class for HHD corpus into four HHD NEs
Declare s: number of HHD corpus sentences
w: HHD word
N: number of unique HHD words
W: word vector of unique words, $W = (w_1, w_2, \ldots w_N)$
M: HHD co-occurrence matrix, generated from |W|
i, j: incremental count variable, initialize as 1
k: array index of word vectors
R_v: preceding context of word vector (row wise)
C_v: followed/succeeding context of word vector (column wise)

Begin:
// Process 1: Generate Word Vector (W)
1. For each pre-processed sentence E
2. W = set(E) // formulate unique word vector
3. EndLoop

// Process 2: Generate HHD Co-Occurance Matrix(M)
4. HAL_model = CountVectorizer(ngram_range = ([1: n])) //generate n-word context window
5. W_Trans = HAL_model.fit_transform(W)
//W_Trans: transformed matrix of W
6. M = ((W_Trans.T) *(W_Trans)) //T: transpose
// Process 3: Similarity Measurement
7. vectors = get_vectors (M, W)
8. LOOP (k<=W, i<=W, j<=W)
9. $R_v[k]$ = vectors (W_{ij}) //construct R_v based on preceding contexts
10. EndLOOP
11. LOOP (k<=W, j<=W, i<=W)
12. $C_v[k]$ = vectors (W_{ij}) //construct C_v based on followed contexts
13. EndLOOP
14. LOOP(k<=W)
15. D[k] = $R_v[k]$ + $C_v[k]$ //construct vectors to compute cosine similarity
16. EndLOOP
17. Cm(D[k_i]D[k_j]) = co_sim(D[k_i], D[k_j])
=dot(D[k_i]D[k_j])/(sqrt(dot(D[k_i]D[k_j])) *sqrt(dot(D[k_j]D[k_j])))
// cosine similarity between two-word vectors
End

Figure 4. HAL Algorithm for Hindi NER

HAL Illustrative Example

Input Hindi Sentences:

"घुटनों के दर्द की सबसे बड़ी वजह है ओवरवेट, जाहिर है आपके भार को सहने में घुटनों को तकलीफ होगी"

(Ghutno ke dard ki sabse badi vajah he overweight, jahir he aapke bhar ko sahane me ghutno ko takleef hogi/ The most common cause of pain in the knees is overweight, obviously will hurt the knee to bear your burden).

The word co-occurrence frequency matrix (M)- a square matrix is generated based upon the number of processed input words. Here, M_{18*18} is constructed as is presented in Table 1. Within the co-occurrence matrix a context window size (n = 5) is chosen to find out the semantic relationships among the words of the matrix. Based on experimental observations, it is seen that if n > 5 then M is highly sparse, words may appear semantically too far from each other. If n < 5 then M is highly dense, moreover, difficult to compute similarities among words, and they may appear too close to each other. Thus, experimental evaluation interprets that the best context size is 5 (n = 5) for words (immediate to distant words, range highest to lowest as 5 to 1) with respect to the HHD corpus.

For an input Hindi word "दर्द" (Dard/Pain), compute preceding context word vector (R_v), succeeding context word vector (C_v) and D[k] as a concatenation of word vector based on R_v and C_v as follows:

$R_v[1]$ = [4 5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]

$C_v[1]$ = [0 0 0 5 4 3 2 1 0 0 0 0 0 0 0 0 0 0]

D[1] = [4 5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5 4 3 2 1 0 0 0 0 0 0 0 0 0 0]

Consider another word "तकलीफ" (Takleef/Problem) Then

$R_v[2]$ = [4 0 0 0 0 0 0 0 0 0 0 0 0 6 2 3 0 0]

$C_v[2]$ = [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5]

D[2] = [4 0 0 0 0 0 0 0 0 0 0 0 0 6 2 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5]

Further, Cosine Similarity (Cm) has been computed between two input words using equation (1)

$$Cm(D[1], D[2]) = \frac{D[1].D[2]}{|D[1]| * |D[2]|} \quad (1)$$

So, the following computations are performed as mentioned below:

D[1]. D[2] = 16

|D[1]| = sqrt((4)^2 + (5)^2 + (5)^2 + (4)^2 + (3)^2 + (2)^2 + (1)^2) = 9.798

Table 1
HAL co-occurrence matrix

	घुटनों	के	दर्द	की	सबसे	बड़ी	वजह	है	ओवरवेट	,	जाहिर	आपके	भार	को	सहने	में	तकलीफ	होगी
घुटनों	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	0	0
के	5	0	0	0	0 0	0	0	0	0	0	0	0	0	0	0	0	0	0
दर्द	4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
की	3	4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
सबसे	2	3	4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
बड़ी	1	2	3	4	5	0	0	0	0	0	0	0	0	0	0	0	0	0
वजह	0	1	2	3	4	5	0	0	0	0	0	0	0	0	0	0	0	0
है	0	0	1	2	3	4	5+1	2	3	4	5	0	0	0	0	0	0	0
ओवरवेट	0	0	0	1	2	3	4	5	0	0	0	0	0	0	0	0	0	0
,	0	0	0	0	1	2	3	4	5	0	0	0	0	0	0	0	0	0
जाहिर	2	0	0	0	0	1	2	3	4	5	0	0	0	0	0	0	0	0
आपके	0	0	0	0	0	0	0	5+1	2	3	4	0	0	0	0	0	0	0
भार	0	0	0	0	0	0	0	4	1	2	3	5	0	0	0	0	0	0
को	5	0	0	0	0	0	0	3	0	1	2	4	5+1	2	3	4	0	0
सहने	0	0	0	0	0	0	0	2	0	0	1	3	4	5	0	0	0	0
में	0	0	0	0	0	0	0	1	0	0	0	2	3	4	5	0	0	0
तकलीफ	4	0	0	0	0	0	0	0	0	0	0	0	0	5+1	2	3	0	0
होगी	3	0	0	0	0	0	0	0	0	0	0	0	0	4	1	2	5	0

|D[2]| = sqrt((4)^2 + (6)^2 + (2)^2 + (3)^2 + (5)^2) = 9.487

Therefore, Cosine Similarity among words "दर्द" (Dard/Pain) and "तकलीफ" (Takleef/Problem) is Cm(D[1], D[2])=0.172.

In addition, cosine similarity among words "दर्द" (Dard/Pain) and "ओवरवेट" (Overweight) can be computed out to be 0.029, while cosine similarity among words "तकलीफ" (Takleef/Problem) and "ओवरवेट" (Overweight) can be computed out to be 0.010. These observations clearly indicate that the two words "दर्द" (Dard/Pain) and "तकलीफ" (Takleef/Problem) are semantically closer to each other, while "दर्द" (Dard/Pain) and "ओवरवेट" (Overweight); "तकलीफ" (Takleef/Problem) and "ओवरवेट" (Overweight) are not semantically close to each other.

Conditional Random Field Framework

Conditional Random Field (CRF) (Li & McCallum, 2003) is a probabilistic based discriminative, undirected graphical model that is highly useable for sequential labelling tasks such as part-of-speech tagging (PVS & Karthik, 2007), table extraction (Pinto et al., 2003), named entity recognition (Ekbal & Bandyopadhyay, 2009), noun phrase segmentation (Sha & Pereira, 2003). CRF has the capability to easily add-on large number of arbitrary, non-independent features in conjunctions to the base features. CRF calculates the conditional probability values on the designated output nodes, given values as are assigned to other designated input nodes.

CRF defines the conditional probability of state sequence $s = <s_1, s_2, s_3...s_T>$, given an observation sequence $o = <o_1, o_2, o_3...o_T>$ as in equation (2):

$$P(s \mid o) = \frac{1}{z_o} exp \sum_{t=1}^{T} \sum_{k=1}^{M} \lambda_k f_k(s_{t-1}, s_t, o, t) \quad (2)$$

Here,

T : number of tokens in a sequence

M : number of features

$f_k(s_{t-1}, s_t, o, t)$: feature function, weight λ_k is learnt via training

z_o : normalization factor over all state sequences

The values of the feature functions may range between $-\infty$, $+\infty$ but typically they are binary. Under binary, $f_k(s_{t-1}, s_t, o, t)$ has value of 0 for most cases, and is only set to be 1 when s_t , s_t are certain states and the observation has certain properties.

Also, to make all conditional probabilities sum up to 1, z_o is defined as in equation (3):

$$z_o = \sum_s \exp(\sum_{t=1}^{T}\sum_{k=1}^{M} \lambda_k f_k(s_{t-1}, s_t, o, t)) \quad (3)$$

In order to train CRF, objective function to be maximized is the penalized log-likelihood of the state sequences, given the observation sequences as in equation (4):

$$L = \sum_{i=1}^{N} \log(P(s^{(i)} \mid o^{(i)}) - \sum_{k=1}^{M} \frac{\lambda_k^2}{2\sigma^2} \quad (4)$$

where,

$\{< o^{(i)}, s^{(i)} >\}$: labelled training data with observed sequence as tokens and state sequence as corresponding labels

$\sum_{k=1}^{M} \frac{\lambda_k^2}{2\sigma^2}$: sum which corresponds to zero-mean

σ^2 : variance, Gaussian prior for parameters optimization

We have used *sklearn-crfsuite 0.3*.6 (https://pypi.python.org/pypi/sklearn-crfsuite) as an open source implementation of CRF for segmenting or labelling sequential HHD corpus.

HHD Feature Engineering

This section details about varied HHD corpus-based features that are used for the experiments.

Head noun feature: Head noun feature (F_{hn}) is usually defined as a major noun or noun phrase of an NE which describes its function or property. It serves as unigram, bigram and trigram head nouns.

Word suffix feature: Word suffix feature (F_{ws}) represents suffix of the current and/or surrounding words. Currently, length of 2 to 4 characters is used as feature. Table 2 shows sample suffixes along with examples from HHD corpus.

Part-of-speech feature: Part-of-speech (POS) feature (F_{ps}) represents the POS information for the previous words and the current word of HHD corpus using POS tagger (https://bitbucket.org/sivareddyg/hindi-part-of-speech-tagger). Several coarse-grained POS tags, such as NomPSP which represents nominal followed by a post-position marker is considered.

N-gram feature: N-gram feature (F_{ng}) extracts n-tuple of HHD corpus words. In this research, only bi-grams and tri-grams are considered while other higher order n-grams are

restricted because of the limitation in the size of the HHD corpus. Table 3 shows some examples of bi-gram and tri-gram features, for all the four HHD corpus-based NE types.

Table 2
Word suffixes and examples

Suffix	HHD Examples
~दर्द	"सिरदर्द", "पेटदर्द", "गलादर्द", "कमरदर्द"
~हट	"अकुलाहट", "मिचलाहट", "झनझनाहट", "सरसराहट", "खरखराहट"
~पन	"दुबलापन", "गंजापन", "चिपचिपापन", "चिड़चिड़ापन", "भारीपन"
~इटिस	"टेन्टीनाइटिस", "बरसाइटिस", "अर्थराइटिस", "आस्टियोअर्थराइटिस"
~त्सक	"चिकित्सक", "मनोचिकित्सक", "दंतचिकित्सक"
~पान	"धूम्रपान", "खानपान"

Table 3
Bi-gram and Tri-gram examples

NE	Bi-gram Examples	Tri-gram Examples
PER	"दंत चिकित्सक" (Dant Chikitsak/Dentist) "वाहन चालक" (Vahan Chalak/Driver)	"हड्डी रोग विशेषज्ञ" (Haddi Rog Visheshagya/ Orthopedic)
DIS	"दमा रोग" (Dama Rog/Asthma) "प्रोस्टेट कैंसर" (Prostate Cancer)	"घुटनों का दर्द" (Ghutno Ka Dard/Knee Pain) "पित्ताशय की पथरी" (Pittashay Ki Pathri/ Gallbladder Stone)
SMP	"खट्टी डकारें" (Khatti Dakare/Belch) "मामूली चोट" (Mamuuli Chott/Minor Injury)	"खट्टी-खट्टी डकारें" (Khatti Khatti Dakare/Belch) "घुटने में सूजन" (Ghutane Me Sujan/Knee Swelling)
CNS	"काले चने" (Kale Chane/Chickpea) "काली मिर्च" (Kali Mirch/Pepper)	"सूखे हरे मटर" (Sukhe Hare Matar/Dry Green Peas) "मूंग की दाल" (Moong Ki Daal/Yellow Lentil)

RESULTS AND DISCUSSION

The proposed system is evaluated using 25K, 50K and 75K HHD blind test corpus; and precision, recall and F-measure metrics are computed. It is observed that as testing goes beyond 75K then there is a stagnation in the performance of the evaluation metric parameters. It so happens because of the occurrence of the overfitting criteria. Overfitting means failing to fit an additional data or fail to reliably predict future observation which arises as the proposed methodology starts memorizing, rather than learning from the HHD

corpus. Table 4 shows the F-measure values for different feature sets in the proposed Hindi NER system. And, to compute F-measure, following categories are considered:

True Positive (TP): system finds NE and is also marked by human annotator.

False Positive (FP): system finds NE but is not marked by human annotator.

True Negative (TN): system does not find NE and is not marked by human annotator.

False Negative (FN): system does not find NE but is marked by human annotator

Hence, precision is the fraction of the correct NE annotations, and is defined as in equation (5):

$$\text{Precision (P)} = \frac{TP}{TP + FP} \tag{5}$$

Recall is the fraction of the NEs that are successfully annotated, and is defined as in equation (6):

$$\text{Recall (R)} = \frac{TP}{TP + FN} \tag{6}$$

F-measure is the weighted harmonic mean of precision and recall, and is defined as in equation (7):

$$\text{F-measure (F)} = \frac{2.P.R}{P + R} \tag{7}$$

While experimenting with various features under 25K, 50K and 75K blind test corpus, it is observed that F_{hn} feature provides lowest F–values for all four NE types as is seen in F1. As features are added such as F_{ws}, F_{ps} to F_{hn} F-values also increases as is seen in F2 and F3 respectively but further adding F_{ng} feature leads to decrease in F-value for SMP NE type on 50K and 75K both, while rest other NEs F-values keeps increasing as seen in F4. When gazetteer lists (F_{gs}) are added to baseline features (F_{hn}, F_{ws}, F_{ps}, F_{ng}) then F-values increase drastically for all four NE types as is seen in F5. HAL is applied for semantic similarity then F_{hl} alone has slight increase in F-values for 25K PER NE but decrease in PER NE for 50K and 75K both as seen in F6. When F_{gs} is accompanied to F_{hl} then again there is a high increase in F-values for all NE types among 25K, 50K and 75K as is seen in F7. F8 and F9 show different combinations of baseline features along with F_{hl} with variations in F-values for different NE types. DIS NE decrease from 84.96 to 84.34 on 50K, CNS NE somewhat increase from 84.04 to 84.57 on 75K, while rest NEs increase in high amount on varied corpus sizes. Finally, F_{hl} along with baseline and F_{gs} give best result for all NE types on 25K, 50K and 75K respectively. F10 shows NEs best F-values, achievable on 75K test as- 90.69%, 89.09%, 87.84% and 88.93% for PER, DIS, SMP and CNS NE types respectively.

Table 4
F-measure for different features in Hindi NER

Feature ID	Feature(s)	F-Measure (%)											
		25K				50K				75K			
		PER	DIS	SMP	CNS	PER	DIS	SMP	CNS	PER	DIS	SMP	CNS
F1	F_{hn}	60.13	61.34	50.04	50.26	62.14	61.76	51.32	51.19	63.12	62.40	52.40	51.37
F2	F_{hn}, F_{ws}	63.12	63.06	54.33	55.11	64.29	64.15	54.78	55.29	65.22	64.56	55.35	55.83
F3	F_{hn}, F_{ws}, F_{ps}	66.55	68.13	62.43	60.15	66.76	69.01	63.66	61.75	67.45	69.33	64.41	62.57
F4	F_{hn}, F_{ws}, F_{ps}, F_{ng}	74.07	76.23	62.55	66.58	74.67	76.92	62.98	67.03	75.76	77.21	63.52	67.35
F5	F_{hn}, F_{ws}, F_{ps}, F_{ng}, F_{gs}	80.18	80.01	72.56	72.36	81.67	80.10	73.10	73.29	82.86	80.13	73.67	74.62
F6	F_{hl}	80.21	81.49	78.04	76.86	80.35	81.56	78.37	77.12	80.79	81.97	78.64	77.78
F7	F_{gs}, F_{hl}	81.39	82.58	80.32	80.01	82.08	83.42	80.64	80.11	82.94	83.84	80.73	80.15
F8	F_{ws}, F_{ps}, F_{ng}, F_{hl}	81.45	83.86	82.45	80.01	81.55	84.96	83.84	80.10	81.52	85.62	84.04	80.01
F9	F_{hn}, F_{ws}, F_{ps}, F_{ng}, F_{hl}	85.45	84.12	84.57	82.16	86.64	84.34	84.57	82.16	87.47	84.28	84.57	82.16
F10	F_{hn}, F_{ws}, F_{ps}, F_{ng}, F_{gs}, F_{hl}	89.08	88.89	86.68	87.13	89.66	88.23	86.38	87.59	90.69	89.09	87.84	88.93

F_{hn}: head noun feature, F_{ws}: word suffix feature, F_{ps}: POS feature, F_{ng}: n-gram feature, F_{gs}: gazetteer lists, F_{hl}: HAL semantic similarity

It is then observed that the overall F-score of different NE types- Person NE, Disease NE, Symptom NE, and Consumable NE for the proposed Hindi NER technique are- 76.98%, 77.42%, 71.57%, and 71.96% respectively which are quite significant as compared with the Maximum Entropy (ME) model (Ahmed & Sathyaraj, 2015: Chieu & Ng, 2002; Curran & Clark, 2003; Hasanuzzaman et al., 2009; Saha et al., 2009, 2008) on the considered Hindi health domain corpus as is seen in Table 5.

Table 5
Comparison of proposed Hindi NER technique w.r.t. Maximum Entropy Model

NE TYPE	Hindi NER Techniques (F-Measure)	
	Proposed Technique	**Maximum Entropy**
PER	76.98%	76.89%
DIS	77.42%	65.34%
SMP	71.57%	53.26%
CNS	71.96%	55.99%

CONCLUSION AND FUTURE SCOPE

In this research work, NER technique for Hindi language using Hyperspace Analogue to Language (HAL) is proposed. HAL uses the semantic based context knowledge which is vital to determine NEs. Such semantics is exploited by the word similarity based on the semantic spaces to cluster words. Four NE types are determined on Hindi health domain (HHD) corpus viz. Person NE, Disease NE, Symptom NE and Consumable NE. Training data passes through HAL steps (word vector generation, co-occurrence matrix, and similarity measurement); feature engineering (head nouns, word suffix, POS and n-gram); manual gazetteer construction and their extension through Hindi WordNet. The training algorithm then estimates parameters for the Conditional Random Field using the trained dataset. Unannotated test data is supplied to the NER system and is transformed into feature vectors for output annotations of the test data. We have used nltk-3.2.4, hal-0.0.3, scikit-learn 0.15.2, sklearn-crfsuite 0.3.6 as set of python modules for the NER task. NEs best F-values, 75K test, achieves 90.69% for Person NE; 89.09% for Disease NE; 87.84% for Symptom NE; 88.93% for Consumable NE respectively. It is observed that the overall F-measure of different NEs on the proposed Hindi NER technique are quite significant as compared to Maximum Entropy model. In future, we intend to focus on the following:

- HHD corpus can be extended to larger extent so that overfitting issue can be resolved to better extent;
- Recognition of some more NE types such as Food, Diagnosis, Treatment;

- Other local semantic techniques such as Correlated Occurrence Analogue to Lexical Semantic (COALS), Random Indexing (RI) can be explored;
- Global context and semantics through Latent Dirichlet Allocation (LDA) can be taken into consideration to enrich word clusters that will lead to better NE accuracy.

REFERENCES

Aggarwal, C. C., & Zhai, C. (Eds.). (2012). *Mining Text Data*. London: Springer Science & Business Media.

Ahmed, I., & Sathyaraj, R. (2015). Named entity recognition by using maximum entropy. *International Journal of Database Theory and Application, 8*(2), 43-50.

Athavale, V., Bharadwaj, S., Pamecha, M., Prabhu, A., & Shrivastava, M. (2016, November). *Towards deep learning in Hindi NER: An approach to tackle the labelled data scarcity*. Retrieved 16 August, 2017, from https://aclweb.org/anthology/W/W16/W16-6320.pdf

Baksa, K., Golović, D., Glavaš, G., & Šnajder, J. (2017). Tagging named entities in Croatian tweets. *Slovenščina 2.0: Empirical, Applied and Interdisciplinary Research*, *4*(1), 20-41.

Baldwin, T., Kim, Y. B., De Marneffe, M. C., Ritter, A., Han, B., & Xu, W. (2015). Shared tasks of the 2015 workshop on noisy user-generated text: Twitter lexical normalization and named entity recognition. In *Proceedings of the Workshop on Noisy User-generated Text* (pp. 126-135). Beijing, China.

Burgess, C., & Lund, K. (1997). Representing abstract words and emotional connotation in a high-dimensional memory space. In *Proceedings of the Cognitive Science Society* (pp. 61-66). Hillsdale, New Jersey.

Chinchor, N., & Robinson, P. (1997, September). MUC-7 named entity task definition. In *Proceedings of the 7th Conference on Message Understanding* (Vol. 29, pp. 1-21). Fairfax, Virginia.

Chiticariu, L., Krishnamurthy, R., Li, Y., Reiss, F., & Vaithyanathan, S. (2010, October). Domain adaptation of rule-based annotators for named-entity recognition tasks. In *Proceedings of the 2010 conference on empirical methods in natural language processing* (pp. 1002-1012). Cambridge, Massachusetts.

Chieu, H. L., & Ng, H. T. (2002, August). Named entity recognition: a maximum entropy approach using global information. In *Proceedings of the 19th international conference on Computational linguistics* (pp. 1-7). Taipei, Taiwan.

Collins, M., & Singer, Y. (1999, June). Unsupervised models for named entity classification. In *1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, *EMNLP* (pp. 100-110). PG County, USA.

Cucerzan, S., & Yarowsky, D. (1999, June). Language independent named entity recognition combining morphological and contextual evidence. In *1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora* (pp. 90-99). PG County, USA.

Curran, J. R., & Clark, S. (2003, May). Language independent NER using a maximum entropy tagger. In *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003* (Vol. 4, pp. 164-167). Edmonton, Canada.

Dey, A., & Prukayastha, B. S. (2013). Named entity recognition using gazetteer method and n-gram technique for an inflectional language: a hybrid approach. *International Journal of Computer Applications*, *84*(9), 31-35.

Ekbal, A., & Bandyopadhyay, S. (2008). A web-based Bengali news corpus for named entity recognition. *Language Resources and Evaluation, 42*(2), 173-182.

Ekbal, A., & Bandyopadhyay, S. (2009). A conditional random field approach for named entity recognition in Bengali and Hindi. *Linguistic Issues in Language Technology*, *2*(1), 1-44.

Ekbal, A., Haque, R., Das, A., Poka, V., & Bandyopadhyay, S. (2008, January). Language Independent Named Entity Recognition in Indian Languages. In *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages* (pp. 33-40). Hyderabad, India.

Ekbal, A., Saha, S., & Sikdar, U. K. (2016). On active annotation for named entity recognition. *International Journal of Machine Learning and Cybernetics*, *7*(4), 623-640.

Farmakiotou, D., Karkaletsis, V., Koutsias, J., Sigletos, G., Spyropoulos, C. D., & Stamatopoulos, P. (2000, September). Rule-based named entity recognition for Greek financial texts. In *Proceedings of the Workshop on Computational lexicography and Multimedia Dictionaries (COMLEX 2000)* (pp. 75-78). Patras, Greece.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. In *Studies in Linguistic Analysis* (pp. 1–32). Blackwell, Oxford.

Grishman, R. (1995, November). The NYU System for MUC-6 or Where's the Syntax. In *Proceedings of the 6th Conference on Message Understanding* (pp. 167-175). Columbia, Maryland, USA.

Gupta, J. P., Tayal, D. K., & Gupta, A. (2011). A TENGRAM method-based part-of-speech tagging of multi-category words in Hindi language. *Expert Systems with Applications*, *38*(12), 15084-15093.

Gupta, V., & Lehal, G. S. (2011). Named entity recognition for Punjabi language text summarization. *International Journal of Computer Applications*, *33*(3), 28-32.

Hasanuzzaman, M., Ekbal, A., & Bandyopadhyay, S. (2009). Maximum entropy approach for named entity recognition in Bengali and Hindi. *International Journal of Recent Trends in Engineering, 1*(1), 408-412.

Jain, A., Yadav, D., & Tayal, D. K. (2014, September). NER for Hindi language using association rules. In *2014 International Conference on Data Mining and Intelligent Computing (ICDMIC)* (pp. 1-5). New Delhi, India.

Jiang, M., Chen, Y., Liu, M., Rosenbloom, S. T., Mani, S., Denny, J. C., & Xu, H. (2011). A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries. *Journal of the American Medical Informatics Association*, *18*(5), 601-606.

Kazama, J. I., & Torisawa, K. (2008). Inducing gazetteers for named entity recognition by large-scale clustering of dependency relations. In *Proceedings of ACL-08: HLT* (pp. 407-415). Columbus, Ohio, USA.

Khalid, M., Jijkoun, V., & De Rijke, M. (2008). The impact of named entity normalization on information retrieval for question answering. In *European Conference on Advances in Information Retrieval* (pp. 705-710). Berlin, Heidelberg: Springer.

Kim, J. H., Kang, I. H., & Choi, K. S. (2002, August). Unsupervised named entity classification models and their ensembles. In *Proceedings of the 19th International Conference on Computational Linguistics-Volume 1* (pp. 1-7). Association for Computational Linguistics, Taipei, Taiwan.

Krishnarao, A. A., Gahlot, H., Srinet, A., & Kushwaha, D. S. (2009, May). A comparison of performance of sequential learning algorithms on the task of named entity recognition for Indian languages. In *Proceedings of the International Conference on Computational Science* (pp. 123-132). Berlin, Heidelberg: Springer.

Konkol, M., Brychcín, T., & Konopík, M. (2015). Latent semantics in named entity recognition. *Expert Systems with Applications*, *42*(7), 3470-3479.

Kozareva, Z. (2006, April). Bootstrapping named entity recognition with automatically generated gazetteer lists. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop* (pp. 15-21). Association for Computational Linguistics, Trento, Italy.

Li, W., & McCallum, A. (2003). Rapid development of Hindi named entity recognition using conditional random fields and feature induction. *ACM Transactions on Asian Language Information Processing (TALIP)*, *2*(3), 290-294.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203-208.

Lund, K., Burgess, C., & Audet, C. (1996). Dissociating semantic and associative word relationships using high-dimensional semantic space. In *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (pp. 603-608). Erlbaum, New Jersey.

Marrero, M., Urbano, J., Sánchez-Cuadrado, S., Morato, J., & Gómez-Berbís, J. M. (2013). Named entity recognition: fallacies, challenges and opportunities. *Computer Standards & Interfaces*, *35*(5), 482-489.

Moro, A., Raganato, A., & Navigli, R. (2014). Entity linking meets word sense disambiguation: a unified approach. *Transactions of the Association for Computational Linguistics*, *2*, 231-244.

Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Lingvisticae Investigationes*, *30*(1), 3-26.

Patil, N., Patil, A. S., & Pawar, B. V. (2016). Survey of named entity recognition systems with respect to Indian and foreign languages. *International Journal of Computer Applications*, *134*(16), 21-26.

Pinto, D., McCallum, A., Wei, X., & Croft, W. B. (2003). Table extraction using conditional random fields. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 235-242). Association for Computing Machinery, Toronto, Canada.

Avinesh, P., & Karthik, G. (2007). Part-of-speech tagging and chunking using conditional random fields and transformation-based learning. In *Proceedings of the IJCAI Workshop on Shallow Parsing for South Asian Languages* (pp. 21-24). Hyderabad, Inda.

Rocktäschel, T., Weidlich, M., & Leser, U. (2012). ChemSpot: a hybrid system for chemical named entity recognition. *Bioinformatics*, *28*(12), 1633-1640.

Rodriquez, K. J., Bryant, M., Blanke, T., & Luszczynska, M. (2012). Comparison of named entity recognition tools for raw OCR text. In *Proceedings of the 11th Conference on Natural Language Processing (KONVENS)* (pp. 410-414). Vienna, Austria.

Saha, S. K., Sarkar, S., & Mitra, P. (2008, January). A Hybrid Feature Set based Maximum Entropy Hindi Named Entity Recognition. In *Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-I* (pp. 343-349). Hyderabad, India.

Saha, S. K., Sarkar, S., & Mitra, P. (2009). Feature selection techniques for maximum entropy based biomedical named entity recognition. *Journal of Biomedical Informatics, 42*(5), 905-911.

Saha, S. K., Mitra, P., & Sarkar, S. (2012). A comparative study on feature reduction approaches in Hindi and Bengali named entity recognition. *Knowledge-Based Systems*, *27*, 322-332.

Saha, S. K., Narayan, S., Sarkar, S., & Mitra, P. (2010). A composite kernel for named entity recognition. *Pattern Recognition Letters*, *31*(12), 1591-1597.

Sahin, H. B., Tirkaz, C., Yildiz, E., Eren, M. T., & Sonmez, O. (2017). Automatically annotated Turkish corpus for named entity recognition and text categorization using large-scale gazetteers. *arXiv preprint arXiv:1702.02363*.

Sha, F., & Pereira, F. (2003, May). Shallow parsing with conditional random fields. In *Proceedings of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1* (pp. 134-141). Association for Computational Linguistics, Edmonton, Canada.

Srivastava, S., Sanglikar, M., & Kothari, D. C. (2011). Named entity recognition system for Hindi language: a hybrid approach. *International Journal of Computational Linguistics (IJCL)*, *2*(1), 10-23.

Szarvas, G., Farkas, R., & Kocsor, A. (2006, October). A multilingual named entity recognition system using boosting and c4. 5 decision tree learning algorithms. In *Proceedings of the International Conference on Discovery Science* (pp. 267-278). Berlin, Heidelberg: Springer.

Tayal, D. K., Ahuja, L., & Chhabra, S. (2015). Word sense disambiguation in Hindi language using hyperspace analogue to language and fuzzy c-means clustering. In *Proceedings of the 12th International Conference on Natural Language Processing* (pp. 49-58). Trivandrum, India.

Zhou, G., & Su, J. (2002). Named entity recognition using an HMM-based chunk tagger. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (pp. 473-480). Association for Computational Linguistics, Philadelphia, Pennsylvania.

Pertanika J. Sci. & Technol. 26 (4): 1823 - 1836 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Statistical Estimators as an Alternative to Standard Deviation in Weighted Euclidean Distance Cluster Analysis

Paul Inuwa Dalatu[1,2]* and Habshah Midi[1,3]

[1]Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2]Department of Mathematics, Faculty of Science, Adamawa State University, Mubi, PMB 25 Mubi Adamawa State, Nigeria
[3]Institute of Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ABSTRACT

Clustering is basically one of the major sources of primary data mining tools. It makes researchers understand the natural grouping of attributes in datasets. Clustering is an unsupervised classification method with the major aim of partitioning, where objects in the same cluster are similar, and objects which belong to different clusters vary significantly, with respect to their attributes. However, the classical Standardized Euclidean distance, which uses standard deviation to down weight maximum points of the *ith* features on the distance clusters, has been criticized by many scholars that the method produces outliers, lack robustness, and has 0% breakdown points. It also has low efficiency in normal distribution. Therefore, to remedy the problem, we suggest two statistical estimators which have 50% breakdown points namely the *Sn* and *Qn* estimators, with 58% and 82% efficiency, respectively. The proposed methods evidently outperformed the existing methods in down weighting the maximum points of the *ith* features in distance-based clustering analysis.

Keywords: Clustering, estimators, K-Means, simulation, weighted

INTRODUCTION

Clustering analysis is an unsupervised learning. It is widely known as unsupervised learning algorithm because it does not involve any statistical assumption to data (Cao et al., 2009). Velmurugan and Santhanam (2011) stated that data modeling places clustering in a historic viewpoint embedded in mathematics, statistics, and

ARTICLE INFO

Article history:
Received: 10 November 2017
Accepted: 23 March 2018
Published: 24 October 2018

E-mail addresses:
dalatup@gmail.com (Paul Inuwa Dalatu)
habshah@upm.edu.my (Habshah Midi)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

numerical analysis. The major aim of clustering is to disintegrate a dataset into dissimilar subsets called clusters or groups, whereby, data in a particular subset have the same membership or characteristics while different subset presenting dissimilar membership from data in distinct subset.

Generally, the current clustering algorithms obtainable in the literature is aimed to offer hard clusters based on K-Means algorithm. The K-Means particularly practices Euclidean distance to measure the alteration between a data object and its cluster centroid. These distances are commonly calculated from raw data and not from normalized data. Whereas, using Euclidean distances, the distance between any two objects is not affected by the addition of new objects to the analysis. The clustering outcomes can be significantly affected by differences in scale among the dimension in which the distances are calculated through. Data normalization is the linear transformation of data to a specific range (Visalakshi &Thangavel, 2009).

Usually, in computing the Euclidean distance function, all features add the same to the function value. Subsequently, different features are usually calculated with different metrics or different magnitudes, and these must be normalized (or standardized) before using the distance function (Munz et al., 2007; Xu & Tian, 2015).

Therefore, one of the issues of K-Means weakness is that it may not always yield global optimum outcomes (Reddy et al., 2012), which necessitates normalizing different metrics when using Euclidean distance function in cluster analysis. However, Xu and Tian (2015) addressed the reported issue and proposed a Standardized Weighted Euclidean Distance:

$$d\left(x_i, x_j\right) = \sqrt{\sum_{i=1}^{n}\left(\frac{x_i - x_j}{s_i}\right)^2}\ , \qquad [1]$$

where s_i (s_i standard deviation of dataset) is an empirical normalization and weighing factor of the *ith* feature. It is observe that the bigger s_i , the smaller is the effect of the *ith* feature on the distance function.

In recent times, some researchers have identified the limitations and drawbacks of the standardized normalization (see Mohamad & Usman, 2013, Jayalakshmi & Santhakumaran, 2011, Jain et al., 2005). Gnanadesikan et al. (1995) studied and conducted experiments on the performance of nine methods on eight most important simulated and real data. Their outcomes revealed and demonstrated weakness of weighting based on the standard deviation. Furthermore, Milligan and Cooper (1988) presented simulation studies on standardization issue. They experimented eight standardization approaches, and the classical Z-Score (i.e. standard deviation) normalization was found to be less effective in various circumstances. However, Sarstedt and Mooi (2014) recommended that in

most cases, normalization by range performed better compared to standard deviation. Furthermore, Matthews (1979) argued that, by down weighting the whole sample using method like Standardized Weighted Euclidean based on variability may probably eliminate the significant between-cluster consistency. Hence, this motivated us to propose *Qn* and *Sn* estimators to replace standard deviation in Standardized Euclidean distance called *Qn* and *Sn* Weighted Euclidean distance, respectively.

This paper is structured as follows: Section 2 provides the materials and methods. Section 3 presents the results and discussion. Section 4 proffers the conclusions of the study.

MATERIALS AND METHODS

Conventional Distance Functions

According to Giancarlo et al. (2010), distance functions are vital components of classification and clustering techniques. Therefore, in comparing performance of the proposed distance function, the K-Means clustering algorithm is also executed using various traditional distance functions, such as the Euclidean and the weighted Euclidean distance.

Euclidean Distance

The most popular distance measure for numerical data is possibly the Euclidean distance, also well-known as L_2 norm, as defined in (Shirkorshidi et al., 2015):

$$d\left(x_i, x_j\right) = \sqrt{\sum_{i=1}^{n}\left(x_i - x_j\right)^2} \qquad [2]$$

This distance measure has the appealing property in which the $d\left(x_i, x_j\right)$ can be interpreted as the physical distance between p-dimensional points $x_i^{`} = x_{i1}, x_{i2}, ..., x_{ip}$ and $x_j^{`} = x_{j1}, x_{j2}, ..., x_{jp}$ in Euclidean space.

Standardized Euclidean

The Standardized Euclidean (or sometimes called Weighted Euclidean) was first proposed by Orloci (1967) based on the fact that Euclidean distance had some demerit of absoluteness on the method. This standardization eliminates the limiting effect of all attribute variables in samples on the maximum likely distance. Therefore, this function will assist in giving equal weight to different values in the set and the distance will become scale invariant. Recently, it was criticized by Gerstenberger and Vogel (2015) that as far as standard deviation was applied to down weight some maximum points, it was prone to outliers and lack robustness. The Standardized Euclidean is computed as in Xu and Tian (2015), Equation [1] revisited.

Proposed Weighted Euclidean Distance Functions

In this section we discuss the two proposed Weighted Euclidean distance functions. The two proposed functions are based on the Weighted Euclidean or sometimes called Standardized Euclidean of Xu and Tian (2015). Xu and Tian (2015) claimed that the larger s_i (denotes the standard deviation of the dataset), the smaller was the influence of the *ith* feature on the distance is. They believed that the rationale behind the method is the assumption that both normal and anomalous may appear from different clusters in the features space. Perhaps, the data may contain outliers which do not belong to a bigger cluster, yet this does not disturb the K-Means clustering as long the number of outliers is small.

Recently, Gerstenberger and Vogel (l2015) criticized the Standardized Weighted Distance in Equation [1] in which it was based on standard deviation to down weight the data. They noted that this method lacked robustness, because the calculation of standard deviation was based on the sample mean which was very sensitive to outliers.

A Standard deviation has 0% breakdown point as stated in Rousseeuw and Hubert (2011). It is susceptible to outliers and has low efficiency in heavy-tailed distributions (Gerstenberger & Vogel, 2015). In order to remedy this problem, we suggest to employing high break down point estimators of *Sn* and Q*n* where both have 50% breakdown points with efficiency of 58% and 82%, respectively (Rousseeuw & Croux, 1993).

The proposed methods are summarized as follows:

***Qn* Weighted Euclidean Distance Function.**

$$d\left(x_i, x_j\right)=\sqrt{\sum_{i=1}^{n}\left(\frac{x_i - x_j}{Qn}\right)^2} \qquad [3]$$

where:

$$Qn = c\left\{\left|x_i - x_j\right|; i < j\right\}, \qquad [4]$$

Rousseeuw and Croux (1993) suggested assigning c = 2.2219 for consistency in the Gaussian distribution.

***Sn* Weighted Euclidean Distance Function.**

$$d\left(x_i, x_j\right)=\sqrt{\sum_{i=1}^{n}\left(\frac{x_i - x_j}{Sn}\right)^2}\ , \qquad [5]$$

where *Sn* is the median of the *n* medians of the absolute difference between x_i and x_j:

$$Sn = c \times med_i \left\{ med_j \left| x_i - x_j \right|; j = 1,2,...,n \right\}, \quad [6]$$

Rousseeuw and Croux (1993) suggested assigning $c = 1.1926$ for consistency in the Gaussian distribution.

RESULTS AND DISCUSSION

Simulation Study

In this section, Monte Carlo simulation study is presented to compare the performance of existing methods, such as Euclidean distance (Loohach & Garg, 2012), and Standardized Weighted Euclidean distance (Xu & Tian, 2015) with our proposed methods *Qn* Weighted Euclidean distance, and *Sn* Weighted Euclidean distance.

Following Loohach and Garg (2012), and Xu and Tian (2015), two (x_1, x_2) and four (x_1, x_2, x_3, x_4) variables are generated such that each of the exploratory variables (x_1, x_2) and (x_1, x_2, x_3, x_4) are simulated from uniform distribution with range [-10, 10]. The variables are clustered into three classes (clusters, groups) as; cluster 1, cluster 2, and cluster 3. We consider a sample of size 50, 100 and 160. The basis for using different sample sizes is to ascertain the consistency, effectiveness and accuracy of the proposed methods compared to the existing methods. The conventional distance functions, Euclidean and Standardized Weighted Euclidean, and the *Qn* and *Sn* Weighted Euclidean Distance Functions were then applied to the data. Some external validity measures such as; purity (Hernandez-Torruco et al. 2014), fowlkes-mallow index (Velardi et al., 2012), rand index (Noorbehbahani et al., 2015: Rand, 1971;), f-measure (score) (Velardi et al., 2012), jaccard index (Velardi et al., 2012), recall (Velardi et al., 2012), f-measure (beta varied) (Velardi et al., 2012), geometric-mean (Tomar & Agarwal, 2015), precision (Kou et al., 2014: Rokach & Maimon, 2008)), specificity (Velardi et al., 2012), accuracy (Tomar & Agarwal, 2015) and sensitivity (Velardi et al., 2012), computing time, and maximum number of clusters clustered are recorded. In each of the experimental runs, there are 1000 replications. The performance of the four methods are evaluated based on average external validity measures for each distance functions, computational timing (minutes), and having three levels of cluster as; cluster 1, cluster 2, and cluster 3. The values in the parenthesis are unnormalized data and not in parenthesis for normalized data. A good method is one that has maximum external validity measure nearly equal to 1 (at maximum 1), less computing time and maximum numbers clustered in each cluster.

Table 1
Average external validity measures, computing time and maximum clusters for n = 50 (x_1, x_2)

Distance Functions	Euclidean	Weighted Euclidean	Qn W. Euclidean	Sn W. Euclidean
Purity	0.866(0.822)	0.869(0.825)	0.893(0.863)	0.893(0.863)
Fow. M. I.	0.867(0.825)	0.876(0.832)	0.892(0.852)	0.892(0.862)
Rand Index	0.895(0.852)	0.894(0.860)	0.930(0.900)	0.931(0.901)
F-M. (Score)	0.854(0.810)	0.857(0.813)	0.882(0.851)	0.881(0.851)
Jaccard Index	0.789(0.745)	0.792(0.768)	0.813(0.790)	0.800(0.793)
Recall	0.866(0.822)	0.867(0.823)	0.897(0.864)	0.891(0.862)
F-M. (varied)	0.852(0.814)	0.866(0.822)	0.887(0.854)	0.885(0.852)
G. Means	0.872(0.837)	0.879(0.846)	0.892(0.865)	0.892(0.867)
Precision	0.878(0.834)	0.879(0.835)	0.897(0.874)	0.896(0.863)
Specificity	0.899(0.855)	0.900(0.890)	0.980(0.953)	0.947(0.914)
Accuracy	0.874(0.831)	0.894(0.850)	0.961(0.930)	0.965(0.932)
Sensitivity	0.866(0.822)	0.867(0.823)	0.898(0.865)	0.899(0.866)
Average	0.865(0.822)	0.870(0.832)	0.902(0.891)	0.907(0.886)
Compt. Time	35(38)	35(37)	28(30)	28(30)
Clust.1(max. 15)	12(11)	11(11)	14(12)	13(12)
Clust.2 (max.15)	10(10)	11(9)	12(11)	12(12)
Clust.3 (max.20)	15(14)	(15)	17(16)	17(16)

Table 2
Average external validity measures, computing time and maximum clusters for n = 50 (x_1, x_2, x_3, x_4)

Distance Functions	Euclidean	Weighted Euclidean	Qn W. Euclidean	Sn W. Euclidean
Purity	0.877(0.844)	0.878(0.845)	0.930(0.901)	0.921(0.916)
Fow. M. I.	0.871(0.846)	0.892(0.863)	0.920(0.911)	0.920(0.911)
Rand Index	0.894(0.861)	0.893(0.860)	0.939(0.922)	0.931(0.924)
F-M. (Score)	0.861(0.832)	0.863(0.830)	0.937(0.914)	0.936(0.903)
Jaccard Index	0.834(0.801)	0.837(0.804)	0.913(0.900)	0.913(0.901)
Recall	0.871(0.843)	0.872(0.845)	0.907(0.890)	0.901(0.890)
F-M. (varied)	0.897(0.864)	0.891(0.865)	0.960(0.930)	0.944(0.911)
G. Means	0.929(0.910)	0.929(0.910)	0.951(0.944)	0.969(0.937)
Precision	0.893(0.864)	0.896(0.864)	0.947(0.914)	0.956(0.923)
Specificity	0.895(0.862)	0.896(0.863)	0.946(0.913)	0.947(0.914)
Accuracy	0.894(0.871)	0.898(0.865)	0.952(0.936)	0.965(0.942)
Sensitivity	0.894(0.871)	0.892(0.870)	0.978(0.945)	0.969(0.936)
Average	0.893(0.855)	0.886(0.857)	0.940(0.920)	0.939(0.917)
Compt. Time	42(45)	42(45)	37(40)	37(40)
Clust.1(max. 15)	11(9)	10(10)	13(13)	13(13)
Clust.2 (max.15)	9(9)	9(9)	13(12)	12(11)
Clust.3 (max.20)	14(13)	15(12)	17(15)	18(16)

Table 3
Average external validity measures, computing time and maximum clusters for n = 100 (x_1, x_2)

Distance Functions	Euclidean	Weighted Euclidean	Qn W. Euclidean	Sn W. Euclidean
Purity	0.894(0.861)	0.894(0.861)	0.960(0.930)	0.963(0.932)
Fow. M. I.	0.895(0.862)	0.895(0.862)	0.950(0.920)	0.950(0.920)
Rand Index	0.897(0.864)	0.898(0.865)	0.970(0.940)	0.964(0.931)
F-M. (Score)	0.896(0.863)	0.897(0.864)	0.967(0.934)	0.969(0.936)
Jaccard Index	0.834(0.801)	0.837(0.804)	0.913(0.900)	0.913(0.900)
Recall	0.894(0.861)	0.895(0.862)	0.937(0.904)	0.937(0.904)
F-M. (varied)	0.897(0.864)	0.895(0.862)	0.933(0.900)	0.937(0.904)
G. Means	0.896(0.863)	0.899(0.866)	0.954(0.921)	0.963(0.930)
Precision	0.895(0.863)	0.959(0.926)	0.970(0.940)	0.970(0.940)
Specificity	0.898(0.865)	0.899(0.866)	0.970(0.940)	0.980(0.950)
Accuracy	0.897(0.864)	0.897(0.864)	0.945(0.912)	0.948(0.915)
Sensitivity	0.894(0.861)	0.894(0.861)	0.960(0.932)	0.960(0.931)
Average	0.891(0.858)	0.897(0.864)	0.952(0.923)	0.955(0.924)
Compt. Time	47(52)	46(52)	41(43)	41(43)
Clust.1(max. 30)	25(24)	26(23)	26(26)	27(25)
Clust.2 (max.30)	20(20)	22(21)	26(23)	26(24)
Clust.3 (max.40)	33(30)	31(30)	35(34)	34(34)

Table 4
Average external validity measures, computing time and maximum clusters for n = 100 (x_1, x_2, x_3, x_4)

Distance Functions	Euclidean	Weighted Euclidean	Qn W. Euclidean	Sn W. Euclidean
Purity	0.896(0.863)	0.899(0.866)	0.950(0.920)	0.954(0.931)
Fow. M. I.	0.896(0.863)	0.898(0.865)	0.953(0.920)	0.953(0.920)
Rand Index	0.899(0.866)	0.898(0.865)	0.970(0.940)	0.964(0.931)
F-M. (Score)	0.894(0.861)	0.893(0.860)	0.970(0.940)	0.969(0.936)
Jaccard Index	0.867(0.834)	0.870(0.840)	0.947(0.914)	0.946(0.913)
Recall	0.896(0.863)	0.894(0.861)	0.941(0.911)	0.935(0.901)
F-M. (varied)	0.893(0.860)	0.893(0.860)	0.933(0.900)	0.937(0.904)
G. Means	0.895(0.862)	0.895(0.863)	0.954(0.0.921)	0.960(0.930)
Precision	0.897(0.864)	0.895(0.862)	0.970(0.940)	0.970(0.941)
Specificity	0.898(0.865)	0.900(0.890)	0.973(0.941)	0.980(0.950)
Accuracy	0.897(0.864)	0.897(0.864)	0.945(0.912)	0.949(0.916)
Sensitivity	0.894(0.861)	0.895(0.863)	0.960(0.930)	0.960(0.930)
Average	0.894(0.861)	0.894(0.861)	0.947(0.924)	0.956(0.925)
Compt. Time	53(55)	53(55)	46(49)	46(49)
Clust.1(max. 30)	23(22)	24(23)	25(24)	26(25)
Clust.2 (max.30)	23(21)	23(20)	25(24)	25(23)
Clust.3 (max.40)	31(28)	30(28)	35(33)	34(33)

Table 5
Average external validity measures, computing time and maximum clusters for n = 160 (x_1, x_2)

Distance Functions	Euclidean	Weighted Euclidean	Qn W. Euclidean	Sn W. Euclidean
Purity	0.911(0.894)	0.914(0.897)	0.963(0.930)	0.963(0.931)
Fow. M. I.	0.912(0.896)	0.912(0.895)	0.963(0.930)	0.953(0.932)
Rand Index	0.940(0.910)	0.948(0.915)	0.970(0.940)	0.965(0.932)
F-M. (Score)	0.910(0.891)	0.913(0.894)	0.970(0.940)	0.970(0.940)
Jaccard Index	0.0.834(0.797)	0.837(0.785)	0.947(0.914)	0.947(0.914)
Recall	0.911(0.894)	0.912(0.892)	0.941(0.911)	0.935(0.902)
F-M. (varied)	0.907(0.900)	0.911(0.891)	0.945(0.914)	0.937(0.910)
G. Means	0.929(0.910)	0.929(0.911)	0.954(0.923)	0.962(0.932)
Precision	0.923(0.890)	0.926(0.913)	0.970(0.941)	0.969(0.936)
Specificity	0.955(0.932)	0.961(0.931)	0.970(0.952)	0.980(0.943)
Accuracy	0.940(0.913)	0.948(0.935)	0.952(0.934)	0.965(0.943)
Sensitivity	0.911(0.893)	0.912(0.890)	0.972(0.941)	0.960(0.934)
Average	0.915(0.893)	0.919(0.896)	0.960(0.931)	0.959(0.932)
Compt. Time	63(64)	63(64)	56(60)	56(60)
Clust.1(max. 50)	46(46)	45(44)	46(46)	46(45)
Clust.2 (max.50)	47(45)	46(44)	47(45)	47(46)
Clust.3 (max.60)	53(53)	54(54)	58(57)	58(58)

Table 6
Average external validity measures, computing time and maximum clusters for n = 160 (x_1, x_2, x_3, x_4)

Distance Functions	Euclidean	Weighted Euclidean	Qn W. Euclidean	Sn W. Euclidean
Purity	0.916(0.912)	0.919(0.915)	0.953(0.943)	0.953(0.943)
Fow. M. I.	0.917(0.915)	0.926(0.922)	0.948(0.932)	0.942(0.932)
Rand Index	0.935(0.922)	0.944(0.930)	0.960(0.948)	0.951(0.940)
F-M. (Score)	0.924(0.911)	0.927(0.913)	0.952(0.945)	0.958(0.945)
Jaccard Index	0.893(0.884)	0.902(0.890)	0.943(0.932)	0.950(0.939)
Recall	0.916(0.902)	0.920(0.913)	0.947(0.934)	0.942(0.936)
F-M. (varied)	0.905(0.891)	0.911(0.892)	0.937(0.924)	0.935(0.922)
G. Means	0.912(0.907)	0.919(0.906)	0.952(0.945)	0.957(0.947)
Precision	0.908(0.894)	0.919(0.895)	0.947(0.934)	0.946(0.936)
Specificity	0.929(0.915)	0.929(0.914)	0.950(0.941)	0.957(0.944)
Accuracy	0.914(0.901)	0.924(0.910)	0.951(0.940)	0.955(0.942)
Sensitivity	0.916(0.902)	0.920(0.913)	0.947(0.934)	0.942(0.936)
Average	0.915(0.905)	0.921(0.909)	0.949(0.938)	0.949(0.939)
Compt. Time	70(72)	70(72)	65(68)	65(68)
Clust.1(max. 50)	45(44)	45(44)	46(45)	47(46)
Clust.2 (max.50)	44(44)	45(43)	47(46)	47(46)
Clust.3 (max.60)	54(53)	55(52)	56(55)	56(55)

Tables 1, 2, 3, 4, 5, and 6 present the average values of 1000 replications of average external validity measures, maximum number of samples in each cluster, and the computing time (minutes).

It is evidently clear that all the two proposed methods have shown impressive performance by achieving the highest average maximum external validity measures and recording lowest computational timing. The proposed methods have also been clustered to nearly the maximum numbers required to be clustered in each cluster (group). This indicates that the performance of the proposed methods is more accurate and efficient compared to the existing methods.

Table 7 presents the average external validity measures and computing time based on 1000 simulation runs, for 5% and 10% contaminated data generated from uniform

Table 7
Average external validity measures and computing time (minutes) for n = 50, 100, and 160

n	Contaminated	Method	x_1, x_2		x_1, x_2, x_3, x_4	
			Av. Ext. Val.	Comp. Time	Av. Ext. Val.	Comp. Time
50	5%	Euclidean	0.6752	53	0.6634	56
		W.ted Eu.	0.6947	52	0.6815	54
		Qn W. Eu.	0.7427	47	0.7141	50
		Sn W. Eu.	0.7503	46	0.7223	49
	10%	Euclidean	0.6178	60	0.5912	61
		W.ted Eu.	0.6331	59	0.6152	62
		Qn W. Eu.	0.6792	54	0.6573	56
		Sn W. Eu.	0.6801	53	0.6633	55
100	5%	Euclidean	0.6973	62	0.6864	64
		W.ted Eu.	0.6978	62	0.6982	65
		Qn W. Eu.	0.7433	59	0.7117	61
		Sn W. Eu.	0.7392	59	0.7114	63
	10%	Euclidean	0.6191	68	0.6043	70
		W.ted Eu.	0.6186	68	0.6128	71
		Qn W. Eu.	0.6707	64	0.6672	66
		Sn W. Eu.	0.6718	64	0.6595	66
160	5%	Euclidean	0.5984	68	0.5873	70
		W.ted Eu.	0.6112	67	0.6004	69
		Qn W. Eu.	0.6673	64	0.6433	67
		Sn W. Eu.	0.6621	63	0.6403	67
	10%	Euclidean	0.5546	69	0.5334	73
		W.ted Eu.	0.5722	70	0.5471	74
		Qn W. Eu.	0.6371	66	0.6013	68
		Sn W. Eu.	0.6494	65	0.6105	67

distribution with range [15, 16]. From the table it is exciting to note that despite the contamination of the data, the two proposed methods; *Qn* weighted Euclidean and *Sn* weighted Euclidean outperformed the two existing methods. Hence, these findings show that the two proposed methods can perform fairly well even in the presence of contamination.

Real Data Applications

In this section, the Iris and Hayes-Roth datasets are considered to verify the performance of our proposed methods:

Iris Dataset. The iris dataset has been used by many researchers, such as Galili (2015), Jayalakshmi and Santhakumaran (2011), Benson-Putnins et al. (2011), and Han et al. (2011). The dataset contains 3 classes of 150 instances each, where each class refers to a type of iris plant. It comprises the following attributes information: (1) Sepal length in cm, (2) Sepal width in cm, (3) Petal length in cm, and (4) Petal width in cm. The classes are listed as follows: (1) iris Setosa, (2) iris Verisiclor, and (3) iris Virginica (Bache and Lichman, 2013).

Hayes-Roth Dataset. The Hayes-Roth dataset has also been used by many researchers, such as Uddin et al. (2017), Han et al. (2011), Jayalakshmi and Santhakumaran, (2011), and Ryu and Eick (2005). The dataset contains 3 classes of 160 instances each, with 4 attributes namely: (1) hobby, (2) age, (3) educational, and (4) marital status (Bache and Lichman, 2013).

The average external validity measures and computing time under each distance function for Iris and Hayes-Roth datasets are presented in Tables 8 and 9, respectively.

The performances of our methods are compared to other methods, and evaluated based on the average external validity measures and computational timing.

It is clear that all the proposed methods have achieved better performance in the two datasets used. It is important to note that Iris data set in Table 8 has recorded the following results; average external validity measures for Euclidean (0.89387), Standardized Weighted Euclidean (0.88002), *Qn* Weighted Euclidean (0.90379) and *Sn* Weighted Euclidean (90262). While, the computing time (minutes) for Euclidean (44), Standardized Weighted Euclidean (44), *Qn* Weighted Euclidean (42) and *Sn* Weighted Euclidean (42). The Hayes-Roth data set in Table 9 has the following results as; external validity measures for Euclidean (0.66190), Standardized Weighted Euclidean (0.66100). *Qn* Weighted Euclidean (0.67031) and *Sn* Weighted Euclidean (0.67166). While, also its computing time (minutes) for Euclidean (45), Standardized Weighted Euclidean (45), *Qn* Weighted Euclidean (43) and *Sn* Weighted Euclidean (43). Generally, on the average, the two datasets indicated that the two proposed methods had shown impressive performance. Therefore, the results based

on the two datasets applied confirmed that the real numbers used in iris dataset provided higher quality performance in the external validity measures compared to integer numbers used in the Hayes-Roth dataset.

Table 8
Average external validity measures and computing time under each Distance Functions for Iris Dataset

Distance Functions	Euclidean	Weighted Euclidean	Qn Weighted Euc.	Sn Weighted Euc.
Purity	0.88667	0.85333	0.89230	0.89107
Fow. M. I.	0.88876	0.85412	0.89117	0.89217
Rand Index	0.92444	0.90222	0.93657	0.93556
F-M. (Score)	0.88609	0.86327	0.89333	0.89111
Jaccard Index	0.80454	0.79793	0.82271	0.82172
Recall	0.88667	0.88533	0.89667	0.89333
F-M. (varied)	0.88528	0.88358	0.89219	0.89229
G. Means	0.91161	0.89848	0.93440	0.93392
Precision	0.89786	0.88599	0.91631	0.91476
Specificity	0.94333	0.93667	0.94743	0.94667
Accuracy	0.92444	0.91222	0.92577	0.92556
Sensitivity	0.88667	0.88533	0.89667	0.89333
Average	0.89387	0.88002	0.90379	0.90262
Compt. Time	44	44	42	42

Table 9
Average external validity measures and computing time under each Distance Functions for Hayes-Roth Dataset

Distance Functions	Euclidean	Weighted Euclidean	Qn Weighted Euc.	Sn Weighted Euc.
Purity	0.61250	0.50375	0.62125	0.62625
Fow. M. I.	0.60666	NaN	0.62372	0.62873
Rand Index	0.77500	0.75002	0.78417	0.77939
F-M. (Score)	0.61236	NaN	0.63216	0.63792
Jaccard Index	0.43881	0.42504	0.44223	0.44135
Recall	0.65132	0.65221	0.65399	0.65333
F-M. (varied)	0.65055	NaN	0.65993	0.66148
G. Means	0.71888	0.70453	0.72183	0.72519
Precision	0.60256	NaN	0.61529	0.61447
Specificity	0.84783	0.83767	0.84861	0.85251
Accuracy	0.77500	0.76250	0.78654	0.78607
Sensitivity	0.65132	0.65221	0.65399	0.65333
Average	0.66190	0.66100	0.67031	0.67166
Compt. Time	45	45	43	43

Note: NaN = Not a Number

CONCLUSION

In this paper, we proposed two methods to overcome the weakness of Standard Weighted Euclidean Distance method, whereby it has 0% breakdown point characteristics (Rousseeuw & Hubert, 2011), a lack of robustness, is susceptible to outliers, and its low efficiency at heavy-tailed distributions (Gerstenberger and Vogel, 2015). The proposed methods are called *Qn* Weighted Euclidean and *Sn* Weighted Euclidean distance functions. These methods are based on the increase of accuracy and efficiency in performance using of high breakdown estimators as *Qn* and *Sn* both have 50% breakdown points, and their efficiency *Sn* is 58% and *Qn* is 82% (Rousseeuw and Croux, 1993). Therefore, to improve the accuracy and efficiency of Standardized Weighted Euclidean (Xu & Tian, 2015), we employed and adopted the ideas of Rousseeuw and Hubert (2011) to make the distribution more symmetric.

Furthermore, we also presented average external validity measures and computing time (minutes) based on 1000 simulation runs for contaminated data. From the outcome, it is exciting to observe that despite the contamination of the data, the two suggested methods had performed better compared to the existing methods. To investigate the performance of our proposed methods, a simulation study and real data were considered. The results indicate that the Euclidean distance function has the least performance. This is due to the fact that the Euclidean distance has not applied any of the existing estimators to down weight the datasets. However, the two proposed methods have good performance; evidently, by achieving nearly maximum points in the average external validity measures, lower computational timing and clustering the object points to almost all their maximum number of cluster centers.

From the results, it can be concluded that the two proposed methods are much better in performance compared to the existing methods.

REFERENCES

Bache, K., & Lichman, M. (2013). *UCI machine learning repository*. Irvine, CA: University of California, School of Information and Computer Science.

Benson-Putnins, D. A. V. I. D., Bonfardin, M., Magnoni, M. E., & Martin, D. (2011). Spectral clustering and visualization: a novel clustering of Fisher's Iris data set. *SIAM Undergraduate Research Online, 4*, 1-15.

Cao, F., Liang, J., & Jiang, G. (2009). An initialization method for the K-Means algorithm using neighborhood model. *Computers & Mathematics with Applications*, *58*(3), 474-483.

Galili, T. (2015). Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, *31*(22), 3718-3720.

Gerstenberger, C., & Vogel, D. (2015). On the efficiency of Gini's mean difference. *Statistical Methods & Applications*, *24*(4), 569-596.

Giancarlo, R., Bosco, G. L., & Pinello, L. (2010, January). Distance functions, clustering algorithms and microarray data analysis. In *International Conference on Learning and Intelligent Optimization* (pp. 125-138). Berlin, Heidelberg: Springer.

Gnanadesikan, R., Kettenring, J. R., & Tsao, S. L. (1995). Weighting and selection of variables for cluster analysis. *Journal of Classification*, *12*(1), 113-136.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques*. *Sabre Foundation.* Waltham, USA: Morgan Kaufmann Publishers.

Hernández-Torruco, J., Canul-Reich, J., Frausto-Solís, J., & Méndez-Castillo, J. J. (2014). Feature selection for better identification of subtypes of Guillain-Barré syndrome. *Computational and Mathematical Methods in Medicine, 2014*, 1-9.

Jayalakshmi, T., & Santhakumaran, A. (2011). Statistical Normalization and Back Propagation for Classification. *International Journal of Computer Theory and Engineering*, *3*(1), 1793-8201.

Kou, G., Peng, Y., & Wang, G. (2014). Evaluation of clustering algorithms for financial risk analysis using MCDM methods. *Information Sciences*, *275*, 1-12.

Liu, W., Liang, Y., Fan, J., Fen, Z., & Cai, Y. (2014). Improved hierarchical K-Means clustering algorithm without iteration based on distance measurement. In *International Conference on Intelligent Information Processing* (pp. 38- 46). Berlin, Heidelberg: Springer.

Loohach, R., & Garg, K. (2012). Effect of distance functions on simple k-means clustering algorithm. *International Journal of Computer Applications*, *49*(6), 7-9.

Ma, M., Luo, Q., Zhou, Y., Chen, X., & Li, L. (2015). An improved animal migration optimization algorithm for clustering analysis. *Discrete Dynamic in Nature and Society, 2015*, 1-12.

Matthews, A. (1979). Standardization of Measures Prior to Cluster-Analysis. B*iometrics, 35*(4), 765-773.

Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification*, *5*(2), 181-204.

Mohamad, I. B., & Usman, D. (2013). Standardization and its effects on K-means clustering algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, *6*(17), 3299-3303.

Münz, G., Li, S., & Carle, G. (2007, September). Traffic anomaly detection using K-eans clustering. In *Proceedings of Performance, Reliability and Dependability Evaluation of Communication Networks and Distributed Systems, GI/ITG Workshop MMBnet* (pp. 13-14). Hamburg, Germany.

Noorbehbahani, F., Mousavi, S. R., & Mirzaei, A. (2015). An incremental mixed data clustering method using a new distance measure. *Soft Computing*, *19*(3), 731-743.

Orloci, L. (1967). An agglomerative method for classification of plant Communities. *The Journal of Ecology, 55*(1), 193-206.

Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, *66*(336), 846-850.

Reddy, D., Jana, P. K., & Member, I. S. (2012). Initialization for k-means clustering using voronoi diagram. *Procedia Technology, 4*, 395-400.

Rokach, L., & Maimon, O. Z. (2008). *Data mining with decision trees: theory and applications* (Vol. 69). Singapore: World scientific Publishing.

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, *88*(424), 1273-1283.

Rousseeuw, P. J., & Hubert, M. (2011). Robust statistics for outlier detection. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *1*(1), 73-79.

Ryu, T. W., & Eick, C. F. (2005). A database clustering methodology and tool. *Information Sciences*, *171*(1), 29-59.

Sarstedt, M., & Mooi, E. (2014). *A concise guide to market research: The Process, Data, and Methods using IBM SPSS Statistics* (pp. 273-318). Berlin Heidelberg: Springer.

Shirkhorshidi, A. S., Aghabozorgi, S., & Wah, T. Y. (2015). A comparison study on similarity and dissimilarity measures in clustering continuous data. *PloS one*, *10*(12), e0144059.

Tomar, D., & Agarwal, S. (2015). Hybrid feature selection based weighted least squares twin support vector machine approach for diagnosing breast cancer, hepatitis, and diabetes. *Advances in Artificial Neural Systems*, *2015*, 1-10.

Uddin, J., Ghazali, R., & Deris, M. M. (2017). An Empirical Analysis of Rough Set Categorical Clustering Techniques. *PloS one*, *12*(1), e0164803.

Velardi, P., Navigli, R., Faralli, S., & Ruiz-Martinez, J. M. (2012, May). A New Method for Evaluating Automatically Learned Terminological Taxonomies. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC* (pp.1498-1504). Istanbul, Turkey.

Velmurugan, T., & Santhanam, T. (2011). An experimental approach. *Information Technology Journal*, *10*(3), 478-484.

Visalakshi, N. K., & Thangavel, K. (2009). Impact of normalization in distributed K-means clustering. *International Journal of Soft Computing*, *4*(4), 168-172.

Xu, D., & Tian, Y. (2015). A comprehensive survey of clustering algorithms. *Annals of Data Science*, *2*(2), 165-193.

Pertanika J. Sci. & Technol. 26 (4): 1837 - 1848 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

An Exploratory Study to Evaluate the Thermal Conductivity of LM25-Borosilicate Glass (P) Composites under the Influence of Different End Chills

Anupama Hiremath[1]* and Joel Hemanth[2]

[1]Department of Mechanical & Manufacturing Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal-576104, Karnataka, India.
[2]School of Engineering, Presidency University, Bengaluru-560064, India

ABSTRACT

The present paper encompasses the fabrication, experimentation, testing and thermal property evaluation of LM25-Borosilicate glass (p) composites obtained through stir casting route with judicious selection and placement of different end chills within the sand molds. The composites required for the present study were cast by melting LM25 aluminum alloy into which varied weight percent of Borosilicate glass powder was introduced under the application of a mechanical stirrer. The melt with required content of reinforcement was introduced in a sand mold containing one end chill. Metallic end chills (copper and mild steel) and non- metallic end chills (silicon carbide and graphite) were employed in the current research. Various thermal tests were conducted on the specimens drawn from near the chill end to evaluate thermal diffusivity and thermal conductivity of the fabricated composite. The analysis of the obtained results illustrate the fact that end chill materials have a pronounced effect on the evaluated thermal properties of the composite as employment of metallic end chills resulted in the reduction of the thermal conductivity of the specimens as opposed with the thermal conductivity values for specimens fabricated with the aid of non-metallic end chills.

Keywords: End chills, stir casting, thermal conductivity

ARTICLE INFO

Article history:
Received: 04 December 2017
Accepted: 25 April 2018
Published: 24 October 2018

E-mail addresses:
anupama.hiremath@manipal.edu (Anupama Hiremath)
joelhemanth@hotmail.com (Joel Hemanth)
* Corresponding author

INTRODUCTION

Recent years have witnessed the development of advanced aluminum alloys that possess excellent strength. But, the introduction of the required quantity and quality of alloying components and secondary processes to bring about the required modification in

ISSN: 0128-7680
e-ISSN: 2231-8526

the final microstructure of the alloy is proven to be a costlier effect yielding a very slight improvement in the strength of the new alloy as compared to the base metal. The ever increasing demand for materials that exhibit immaculate strength to weight ratio coupled with superior mechanical and thermal properties has led to greater innovation in the field of aluminum metal matrix composites reinforced with nano-to- micro meter scale particulate reinforcements (Mousavian et al., 2016, pp. 58-70; Ma et al., 2014, pp. 366-373). Particle reinforced aluminum matrix composites (PAMCs) (Bodunrin et al., 2015), are known to demonstrate considerable advancement in tensile strength, resistance to wear and also exhibit excellent tailorability of the physical and thermal properties such as coefficient of thermal expansion, density and many more. Due to these derived advantages particle reinforced aluminum matrix composites are preferred over un-reinforced composites for practical applications (Kumar et al., 2016). PAMCs are composed of aluminum and its alloy as matrix reinforced with oxide, boride and carbide ceramic particles. The soft aluminum matrix (Ramnath et al., 2014; Nestler et al., 2014, pp. 125-130) imparts the required ductility, formability, conductivity of heat and electricity. On the other hand, the hard ceramic particles infuse the composite with excellent hardness, strength, resistance to wear and heat (Ma et al., 2017). PAMCs are fast replacing conventional monolithic metals in automotive applications which constantly demand for an extremely low strength- to-weight ratio to gain an excellent fuel economy (Saito et al., 2000). This calls for a thorough study of composite properties at elevated temperatures (200°C-550°C) to validate its use in automotive applications (Ravi Kumar et al., 1999, Garb et al., 2017). Thermal property evaluation is imperative for successful implementation of PAMCs in various automotive, aerospace, electronic packaging and various other heat management applications (Krishna et al., 2016). An effort has been made in the present paper to investigate thermal diffusivity and thermal conductivity of the fabricated PAMC to analyze its behavior at higher operating temperatures. Researchers are experimenting with glass, a most widely used ceramic material, for incorporation as reinforcement within aluminum and other metal matrices. Glass is easily available in all commercial grades at a very low cost and is known for its hardness and excellent resistance to very high operating temperatures. Review of the available literature also proves the fact that glass reinforced aluminum matrix composites possess appreciable mechanical and thermal properties over their matrix metal (Khoramkhorshid et al., 2016; Gopal et al., 2017). But, less work has been done to study the effect of borosilicate glass particulate reinforcement on the properties of aluminum and its alloys. An effort is made in the current research work to analyze the effect of borosilicate glass addition into LM25 matrix on the thermal conductivity of the PAMCs. Aluminum alloys pose a difficulty in feeding them into the mold cavity as they tend to cool over a varied range of temperature through the entire course of solidification. Thus it becomes

imperative to strictly control various process parameters of casting. Researchers have found that judicious inclusion of end chills in the mold cavity results in drastic improvement in the quality of the PAMC castings (Hiremath et al., 2017).

MATERIALS AND METHODS

LM25 as matrix material

LM25 aluminum alloy is a cast alloy which finds its extensive application in the field of automobile industries where it is used to fabricate components such as cylinder heads and cylinder blocks that call for a material with excellent resistance offered against wear, corrosion and can retain its structural stability even at high operating temperatures (Bhatija et al., 2017). In the current research work, LM25 was selected as the matrix material due to its ability to be casted easily into any intricate shape with excellent accuracy. The elemental composition on weight percent basis of LM25 is illustrated in Table 1.

Table 1
Elemental composition of LM25

Elements	Zinc	Magnesium	Silicon	Copper	Manganese	Iron	Aluminum
Composition (Wt. %)	0.10	0.37	7.00	0.20	0.10	0.20	Balance

Borosilicate Glass Particles as Reinforcements

The work investigates the employability of the cast composites in severe operating temperatures which calls for the incorporation of thermally stable reinforcements. Hence, borosilicate glass, which is thermally very stable, was used as reinforcement. The available literature illustrates the fact that glass can be effectively used as reinforcement in PAMCs to obtain a newer composite exhibiting superior strength and structural stability (Zhang et al., 2017). In the current work borosilicate glass particles of 100 μm size was incorporated within the LM25 matrix as reinforcement. Borosilicate glass is a particular type of glass in which silica and boron trioxide elements are the main glass-forming constituents. Borosilicate glass possesses superior mechanical and thermal properties coupled with excellent resistance to wear and corrosion and possessing a very low coefficient of thermal expansion and hence, making it suitable to be employed in severe operating conditions (Hiremath et al., 2017). The elemental composition of borosilicate glass on weight percent basis, is approximately 80% silica, 13% boric oxide, 4% sodium oxide and 2–3% aluminum oxide. Table 2 depicts the mechanical and thermal properties of borosilicate glass.

Table 2
Mechanical and thermal properties of Borosilicate Glass

Mechanical Properties	
Compressive strength	2000 MPa
Young's Modulus	68 GPa
Ultimate Tensile Strength	280 MPa
Knoop Hardness	510
Thermal Properties	
Specific Heat Capacity	830 J/kgK
Coefficient of Thermal expansion	3.3 µm/m-K
Thermal Conductivity	1.2 W/mK

Materials for End Chills

The employment of end chills together with stir casting is made to obtain superior quality castings with minimum to zero defects through the promotion of controlled directional solidification. This is achieved by the ability of the chill to extract heat at a faster rate from the melt at the chill-composite interface, thus promoting directional solidification of the entire casting (Hemanth, 2017). External end chills possessing different volumetric heat capacities (VHC) were utilized in the current research to obtain sound castings with superior mechanical and thermal properties such as tensile strength, hardness and thermal conductivity. Table 3 depicts the thermal properties of the end chills employed in the current research.

Table 3
Thermal properties of the end chills

End chill material	Density (g/cm^3)	Specific heat (J/kg K)	Coefficient of Thermal Expansion ($\alpha \times 10^{-6}$/°C)	Thermal conductivity (W/m K)
Silicon carbide	3.21	650	4.0	120
Graphite	1.95	710	7.1	150
Copper	8.96	385	16.2	385
Mild Steel	7.85	502	12	66

Chill Casting

End chills of the required dimensions were carefully placed in the thoroughly air-dried sand mold cavity before the introduction of the melt (molten alloy mixed with required quantity of reinforcement) into the cavity. Two metallic chills made of copper and mild steel and two non-metallic chills made of silicon carbide and graphite were selected for the present research. The molds for the rectangular castings were prepared using silica sand with 5% bentonite and 5% moisture. The mold dimensions and arrangement of end chills in the mold cavity is as shown in Figure 1.

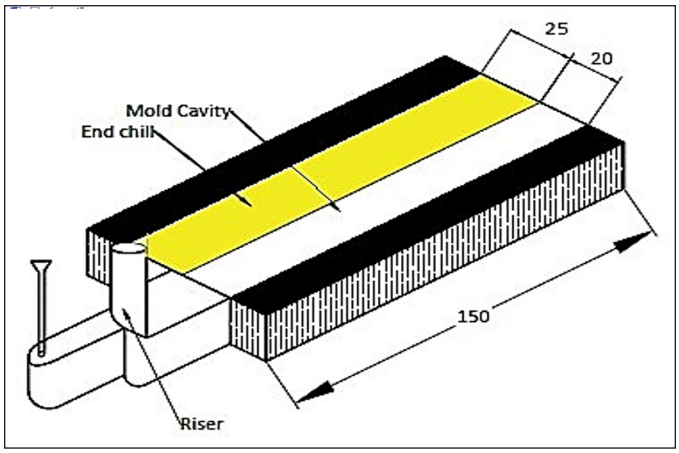


Figure 1. Placement of the End Chills in the Mold Cavity

Vortex Technique for Casting Particle Reinforced Composites

Vortex technique which is also called as stir casting is the most viable commercial technique for fabricating particle reinforced composites due to its numerous advantages that include low-cost, tailorability, simplicity and ease of adaptation in mass production scenarios. The vortex technique involves the addition of particulate reinforcements into molten metal matrix while the melt is simultaneously stirred with the help of a mechanical stirrer and thus producing a vortex that promotes uniform distribution of the particles within the melt (Singh et al., 2017). The LM25 alloy was melted in an electric furnace at approximately 750˚C; preheated borosilicate glass particulates (100 micrometers) were introduced in varied weight percent starting from 3 weight percent and going up till 12 weight percent with an increment of 3 weight percent in consecutive steps. The molten melt was stirred with the help of a mechanical stirrer rotated at 800 rpm for about 5 minutes. The melt was then introduced in properly air-dried green sand molds containing end chills. Figure 2 illustrates the solidified castings in the mold cavity.

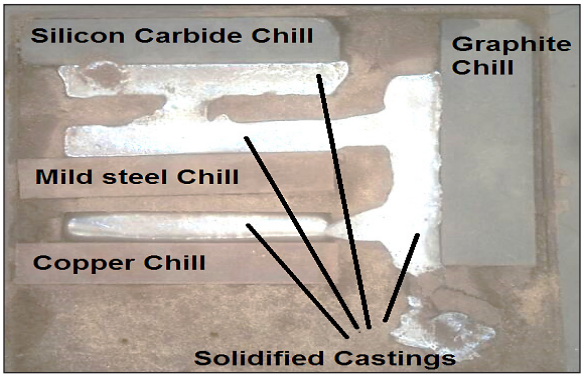


Figure 2. Solidified castings in the mold cavity

Density Measurement

The theoretical density (ρ) of the fabricated composite is obtained from the rule of mixtures as follows:

$$\rho_C = (V_M \times \rho_M) + (V_R \times \rho_R) \quad (1)$$

Where ρ is the density, V is the volume and suffixes C, M and R represent the cast LM25-borosilicate glass composite, LM25 matrix and borosilicate glass reinforcement respectively. The density of LM25 alloy used in the present work was 2.68 g/cc and the density of borosilicate glass particulate reinforcement was 2.23 g/cc. Experimental density was calculated by using the Archimedean water displacement method at ambient conditions.

Thermal Conductivity Test

The determination of thermal conductivity is achieved by the laser flash technique. The NETZCH LFA 447 Nano Flash ® equipment was used in the present work to evaluate the thermal conductivity of the specimens. Thermal conductivity of the sample can be calculated by measuring its thermal diffusivity provided its specific heat and density are known. The specific heat capacity of the composites was determined with the help of NETZCH DSC 200 f3 Maia differential scanning calorimeter equipment. If the density of the sample is known then thermal conductivity of that sample is obtained through:

$$\lambda_T = a_T \times C_{p_T} \times \rho_T \quad (2)$$

Where λ_T the thermal conductivity of the test sample, a_T the thermal diffusivity of the test sample, C_{p_T} the specific heat capacity and ρ_T the density of the test sample. The square samples of dimensions 10 × 10 × 3 mm were machined near the chill end of the Metal Matrix Composites (MMCs). The specimens were coded for recognition and were tested from room temperature to 300ºC. The specimens were coated with graphite spray on their front and rear side to improve the absorptivity of the flash on the front side and emissivity from the rear side.

RESULTS AND DISCUSSION

Density of the Composites

Table 4 illustrates the theoretical and experimental density of all the specimens with varied reinforcement content calculated using the above stated methods. Figure 3 illustrates the percentage difference in density for the cast MMCs.

Figure 4 depicts the microstructure of the composites reinforced with 6 wt.% reinforcement cast with the aid of metallic and non-metallic end chills. It is clear from the

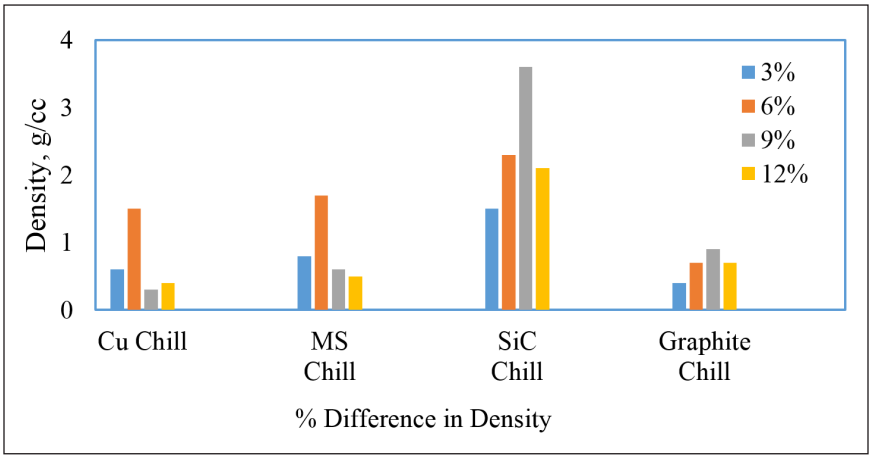
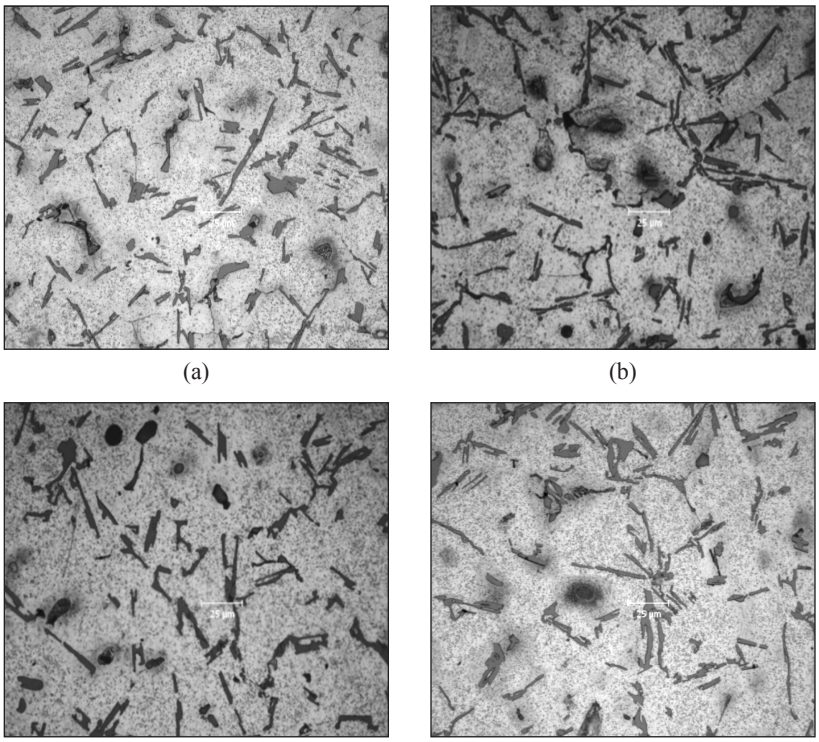


Figure 3. Percentage difference in density

Table 4
Specific heat capacity, experimental and theoretical densities of cast composites

End chills	LM25-Borosilicate glass Composites	Density in g/cc		Error	Specific heat capacity, Cp of cast MMCs (J/kgK)
	Wt.% Reinforcement	Theoretical	Experimental		
Copper	3%	2.665	2.671	0.006	950
	6%	2.653	2.668	0.015	947
	9%	2.639	2.642	0.003	940
	12%	2.626	2.630	0.004	938
Mild steel	3%	2.665	2.673	0.008	951
	6%	2.653	2.670	0.017	946
	9%	2.639	2.645	0.006	940
	12%	2.626	2.631	0.005	939
Silicon carbide	3%	2.665	2.680	0.015	950
	6%	2.653	2.676	0.023	945
	9%	2.639	2.675	0.036	941
	12%	2.626	2.647	0.021	940
Graphite	3%	2.665	2.669	0.004	952
	6%	2.653	2.660	0.007	944
	9%	2.639	2.648	0.009	940
	12%	2.626	2.633	0.007	939

microstructural examinations that the composites cast with the help of metallic end chills possess a refined grain structure. This is due to the fact that metallic end chills possess higher volumetric heat capacities and hence promote directional solidification of the melt (Hiremath et al., 2017; Hemanth, 2009; Anantha Prasad et al., 2015). Thus, composites cast with silicon carbide end chill possess a greater difference in their experimental and theoretical density values.



(a) (b)

(c) (d)

Figure 4. Microstructure of Composites (6 wt.% reinforcement) Cast using Metallic and Non-metallic end chills. (a) Copper chill (b) Mild steel chill (c) Silicon carbide chill (d) Graphite chill

Specific Heat Capacities of the Composites

The specific heat capacities of the specimens are measured with the aid of NETZCH DSC 200 f3 Maia differential scanning calorimeter equipment and the results are tabulated. Table 4 illustrates the specific heat capacity values for the cast MMC samples.

Thermal Conductivity of the Composites

Figures 5, 6, 7 and 8 illustrate the variation of thermal conductivity of the fabricated composites with respect to temperature.

Analysis of the above mentioned result indicates that the thermal conductivity of the fabricated composites reduce with temperature and also with an increase in the borosilicate glass particulate content within the matrix. This is because borosilicate glass particulate reinforcement has an excellent resistance to temperature and a very low thermal conductivity of 1.2 W/m K at 90°C. This acts as a thermal insulator and thus lowers the thermal conductivity of the fabricated composite material. As the percent content of borosilicate glass particulate increases within the matrix, the thermal conductivity of the

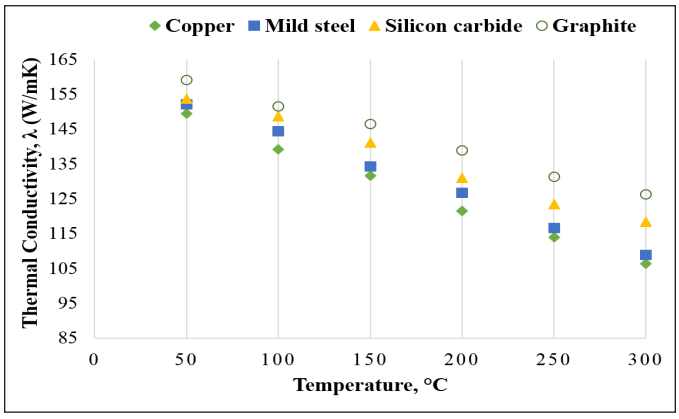


Figure 5. Thermal conductivity vs temperature for LM25-3 wt. % borosilicate glass particulate composite fabricated with different end chill

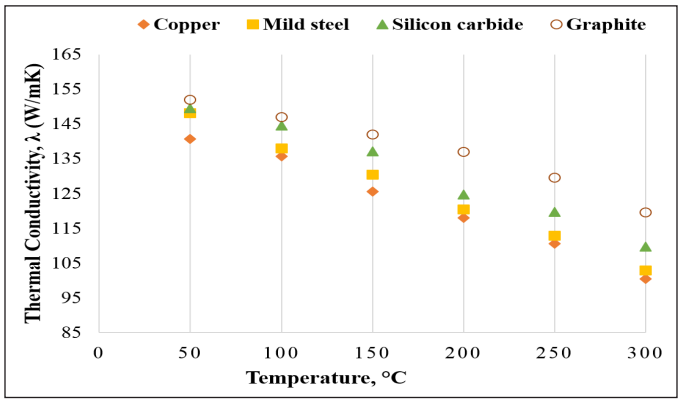


Figure 6. Thermal conductivity vs temperature for LM25-6 wt. % borosilicate glass particulate composite fabricated with different end chill

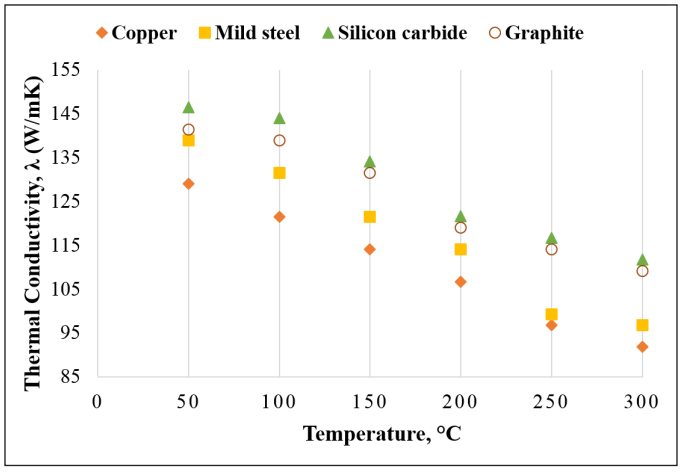


Figure 7. Thermal conductivity vs temperature for LM25-9 wt. % borosilicate glass particulate composite fabricated with different end chill

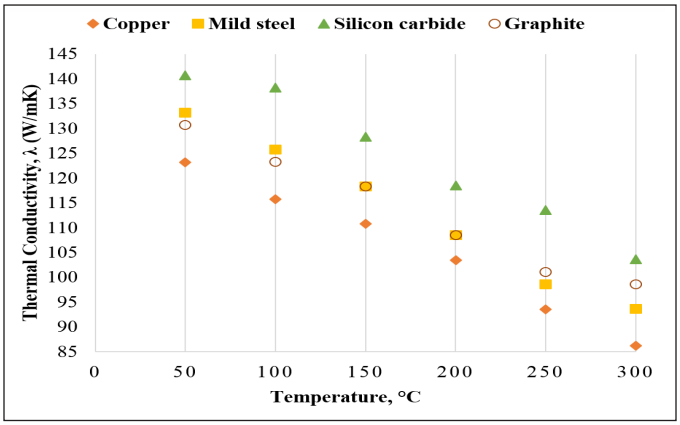


Figure 8. Thermal conductivity vs temperature for LM25-12 wt. % borosilicate glass particulate composite fabricated with different end chill

MMC also reduces. The thermal conductivity values for MMCs are less than the thermal conductivity value of LM25 (170 W/m K at 300°C). Analyzing the above mentioned thermal conductivity values we can conclude that the composite fabricated with the help of copper chill has the least value in comparison to the composite samples fabricated with the help of mild steel, silicon carbide and graphite end chills. Because copper chill with its higher VHC ensures a uniform distribution of borosilicate glass particles within the matrix compared to other end chills. The uniform distribution of the thermally resistant reinforcement results in the overall reduction of the thermal conductivity of the MMC.

CONCLUSIONS

The following conclusions can be drawn from the aforesaid statements:

- Increasing the borosilicate glass particulate content within the matrix is not beneficial for the thermal conductivity of the composite material.
- Chilling has a pronounced effect on the quality of the castings obtained and thus MMCs produced by the employment of metallic end chills have experimental densities almost equal to theoretical densities.
- The thermal conductivity values are least for composites fabricated with copper end chill followed by mild steel, silicon carbide and graphite end chills respectively.

REFERENCES

Bhatija, K. A., & Radhika, N. (2017). Studies on sliding wear characteristics of aluminium LM25/silicon dioxide functionally graded composite and optimisation of parameters using response surface methodology. *Materialwissenschaft und Werkstofftechnik, 48*(6), 600-610.

Bodunrin, M. O., Alaneme, K. K., & Chown, L. H. (2015). Aluminium matrix hybrid composites: A review of reinforcement philosophies; mechanical, corrosion and tribological characteristics. *Journal of Materials Research and Technology, 4*(4), 434-445.

Garb, C., Leitner, M., & Grün, F. (2017). Effect of elevated temperature on the fatigue strength of casted AlSi8Cu3 aluminium alloys. *Procedia Structural Integrity, 7*, 497-504.

Gopal, P. M., Prakash, K. S., Nagaraja, S., & Aravinth, K. (2017). Effect of weight fraction and particle size of CRT glass on the tribological behaviour of Mg-CRT-BN hybrid composites. *Tribology International, 116*, 338-350.

Hemanth, J. (2009). Quartz (SiO_2p) reinforced chilled metal matrix composite (CMMC) for automotive applications. *Materials & Design, 30*(2), 323-329.

Hemanth, J. (2017). Effect of Metallic, Nonmetallic, Water Cooled and Cryogenic Chills on Pearlite Content (PC), Eutectic Cell Count (ECC) and Grain Size (GS) of Hypo Eutectic Nickel Alloyed Cast Iron. *Modeling and Numerical Simulation of Material Science, 7*(1), 1.

Hiremath, A., & Hemanth, J. (2017). Experimental Evaluation of the Chill Casting Method for the Fabrication of LM-25 Aluminum Alloy-Borosilicate Glass (p) Composites. In *6th International Conference of Advanced Materials and Engineering Materials, ICAMEM* (Vol. 748, pp. 69-73). Singapore.

Hiremath, A., & Hemanth, J. (2017). Experimental Evaluation of the Coefficient of Thermal Expansion of Chilled Aluminum Alloy-Borosilicate Glass (p) Composite. *Journal of Materials and Environmental Sciences, 8*(12), 4246-4252.

Khoramkhorshid, S., Alizadeh, M., Taghvaei, A. H., & Scudino, S. (2016). Microstructure and mechanical properties of Al-based metal matrix composites reinforced with Al 84 Gd 6 Ni 7 Co 3 glassy particles produced by accumulative roll bonding. *Materials & Design, 90*, 137-144.

Krishna S. A. M., Shridhar T. N., & Krishnamurthy L. (2015). Microstructural Characterization and Thermal Conductivity Behaviour of Al-6061-SiC-Gr Hybrid Metal Matrix Composites. *Indian Journal of Engineering & Materials Sciences, 23*(4), 207-222.

Kumar, A., Kumar, M., & Goyal, R. (2016). Mechanical behavior of Particle Reinforced Aluminium Matrix Composites. *International Journal for Innovative Research in Science & Technology, 2*(12), 436-439.

Ma, K., Lavernia, E. J., & Schoenung, J. M. (2017). Particulate Reinforced Aluminum Alloy Matrix Composites–A Review on the Effect of Microconstituents. *Reviews on Advanced Materials Science, 48*(2), 91-104.

Ma, Z., Samuel, A. M., Doty, H. W., Valtierra, S., & Samuel, F. H. (2014). Effect of Fe content on the fracture behaviour of Al–Si–Cu cast alloys. *Materials & Design, 57*, 366-373.

Mousavian, R. T., Khosroshahi, R. A., Yazdani, S., Brabazon, D., & Boostani, A. F. (2016). Fabrication of aluminum matrix composites reinforced with nano-to micrometer-sized SiC particles. *Materials & Design, 89*, 58-70.

Nestler, A., & Schubert, A. (2014). Surface properties in ultrasonic vibration assisted turning of particle reinforced aluminium matrix composites. *Procedia CIRP, 13*, 125-130.

Prasad, M. A., & Bandekar, N. (2015). Study of microstructure and mechanical behavior of aluminum/garnet/carbon hybrid metal matrix composites (HMMCs) fabricated by chill casting method. *Journal of Materials Science and Chemical Engineering, 3*(03), 1-8.

Ramnath, B. V., Elanchezhian, C., Jaivignesh, M., Rajesh, S., Parswajinan, C., & Ghias, A. S. A. (2014). Evaluation of mechanical properties of aluminium alloy–alumina–boron carbide metal matrix composites. *Materials & Design, 58*, 332-338.

Ravi Kumar, D., & Swaminathan, K. (1999). Tensile deformation behaviour of two aluminium alloys at elevated temperatures. *Materials at high temperatures, 16*(4), 161-172.

Saito, M., Iwatsuki, S., Yasunaga, K., & Andoh, K. (2000). Development of aluminum body for the most fuel efficient vehicle. *JSAE Review, 21*(4), 511-516.

Singh, R., Singh, R., Dureja, J. S., Farina, I., & Fabbrocino, F. (2017). Investigations for dimensional accuracy of Al alloy/Al-MMC developed by combining stir casting and ABS replica based investment casting. *Composites Part B: Engineering, 115*, 203-208.

Zhang, L., Yang, L., Leng, J., Wang, T., & Wang, Y. (2017). Alloying Behavior and Properties of Al-Based Composites Reinforced with Al85Fe15 Metallic Glass Particles Fabricated by Mechanical Alloying and Hot Pressing Consolidation. *JOM, 69*(4), 748-755.

Pertanika J. Sci. & Technol. 26 (4): 1849 - 1866 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Performance Analysis of Isolated Speech Recognition System Using *Kannada* Speech Database

Ananthakrishna Thalengala[1]*, Kumara Shama[2] and Maithri Mangalore[3]

[1,2]Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal, India
[3]Department of Mechatronics, Manipal Institute of Technology, Manipal, India

ABSTRACT

In this article, performance analysis of speech recognition system for different acoustical models has been presented. In the present work, one of the well-known south Indian language named "*Kannada*" language is considered. Significantly large amount of work has been reported for Automatic Speech Recognition (ASR) in European languages whereas quite a small number of publications can be found in Indian languages. One of the reasons for this gap is that standard speech database in Indian languages is not available. In this study, *Kannada* speech corpus based on *Kannada* broadcast news data has been developed. The isolated speaker independent speech recognition system has been developed using Hidden Markov Tool Kit (HTK). The system front-end uses Mel frequency cepstral coefficients (MFCC) and its derivatives as acoustic features whereas acoustical models are developed by using Hidden Markov Models (HMM). Syllable and mono-phone based *Kannada* dictionaries have been developed in this study. Various mono-phone models considered in this work are word-level, syllable-level and phone-level models. Further, performance evaluation of mono-phone and tri-phone acoustical models for large sized dictionary also carried out. The best word recognition accuracies of 67.82% and 70.56% are reported for mono-phone and tri-phone based systems respectively. The recognition results for different HMM based acoustical models are obtained and hence the recognition performance has been analyzed.

ARTICLE INFO

Article history:
Received: 04 July 2016
Accepted: 17 April 2018
Published: 24 October 2018

E-mail addresses:
anantha.kt@manipal.edu (Ananthakrishna Thalengala)
shama.kumar@manipal.edu (Kumara Shama)
maithri.m@manipal.edu (Maithri Mangalore)
* Corresponding author

Keywords: Hidden Markov Tool Kit (HTK), *Kannada* language, Mel frequency cepstral coefficients (MFCC), Isolated Word Recognition (IWR) system, mono-phone model, phone dictionary, syllable dictionary, tri-phone model

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

One of the most natural and powerful means of human to human interactions is through speech communications. The human auditory perception system possesses natural ability to recognize speech with highest accuracy regardless of speech variations such as environmental noise, speaker's characteristics, emotions of the speaker, speaking rate and so on. And yet another speech variability is speaking accent which is based on the language or region. Even though there are significantly large number of research developments in ASR and state of the art speech recognition systems available, the system performances are not anywhere near to human auditory perception system. This is mainly because of the fact that speech is characterized by multi-modal variability such as emotion, speaker's characteristics, co-articulation effect, accent, and background noise. Among others, co-articulation effect and accent are more challenging variability to be addressed. It would be more appropriate to develop speech recognition system to specific application than looking for a generic applications. In this paper *Kannada* speech is considered with its applications in automatic name dialing (small vocabulary), railway information retrieval system (medium vocabulary) and broadcast news transcriptions (large vocabulary). A few of research contributions related to speech recognition for Indian languages are in Hindi (Kumar, Rajput, & Verma, 2004; Kumar, Aggarwal, & Jain, 2012; Saini, Kaur, & Dua, 2013), Tamil (Lakshmi & Murthy, 2006; Thangarajan, Natarajan, & Selvam, 2009; Radha, 2012), Telugu (Sunitha & Kalyani, 2012; Vijai Bhaskar, Rao, & Gopi, 2012), Marathi (Gawali, Gaikwad, Yannawar, & Mehrotra, 2011), Assamese (Bharali & Kalita, 2015), Punjabi (Dua, Aggarwal, & Kadyan, 2012) and *Kannada* (Hegde, Achary, & Shetty, 2012) (Punitha & Hemakumar, 2014; Hegde, Achary, & Shetty, 2015; Shridhara, Banahatti, Narthan, Karjigi, & Kumaraswamy, 2013).

LITERATURE REVIEW

In the recent few years, research in speech recognition for Indian languages is getting accelerated. Some of the research articles related to Hindi speech recognition are by Kumar, Rajput and Verma (2004), Kumar, Aggarwal and Jain (2012), and Saini, Kaur and Dua (2013). In these articles, authors have explored the speech recognition performance based on word-level, phoneme-level and tri-phone modelling methods. In most of the works authors have considered small or medium sized dictionary. The large vocabulary continuous speech recognition (LVCSR) system for Hindi reported a word recognition accuracy of 75% by using tri-gram language model (Kumar, Rajput, & Verma, 2004). A syllable based continuous speech recognition for Tamil language has been developed and shows that the accuracy of the proposed system is comparable with that of the baseline tri-phone system

(Thangarajan, Natarajan, & Selvam, 2009). A small vocabulary isolated word recognition system in Tamil has been reported a word recognition accuracy of 88% (Radha, 2012). Another work on syllable based speech recognition system for Telugu has reported speech recognition accuracy of 80% (Sunitha & Kalyani, 2012). Recently, Bharali and Kalita (2015) built an isolated word recognition system for Assamese language with vocabulary size of ten words and had reported maximum word recognition accuracy of 95%. In this work, performances of various speech parameters such as linear predictive coding (LPC) analysis, LPC cepstral coefficients (LPCEPSTRA), and Mel frequency cepstral coefficients (MFCC) have been explored. The 39-length MFCC (along with its derivatives) features have shown the best recognition accuracy. Panda & Nayak,(2016) had come out with new syllable segmentation technique for Indian languages, which would help to improve performance of syllable centric speech recognition systems. In this article authors have considered speech samples from three Indian languages, viz. Hindi, Odia, and Bengali.

The literature for speech recognition systems based on *Kannada* language speech corpus is very much limited. Hegde, Achary and Shetty (2012) had developed isolated speech recognition system for *Kannada* words using support vector machine (SVM) technique. The objective of the work was to explore SVM classifier for small vocabulary system. The SVM had been trained by using MFCC parameters of the speech and reported a word recognition accuracy of about 79% for a small vocabulary size of 10 *Kannada* words. Further, Hegde, Achary and Shetty, (2015) had extended their work towards the classification of 'alpha-syllabary' sounds in *Kannada* language. Alpha-syllabary units basically represent alphabets ('*Aksharas*') of *Kannada* language. Statistical analysis of acoustical features such as MFCC and LPC had been carried out and classification accuracies of individual vowels and consonants of *Kannada* using SVM and HMM classifiers had been explored. In another work by Shridhara, Banahatti, Narthan, Karjigi and Kumaraswamy (2013), a prosodically guided phonetic search engine for *Kannada* speech corpus had been implemented. Speech utterances in three different contexts namely, read mode, conversation mode, and extempore mode were considered in this work. The transcription of *Kannada* speech had been carried out in different layers and conventional speech recognition system using HTK had been built and phone recognition accuracies were analyzed. In our earlier study (Ananthakrishna, Maithri, & Shama, 2015; Thalengala & Shama, 2016), isolated speech recognition system based on *Kannada* speech corpus was presented. Comparative study of syllable based and phone based acoustical models were made for a medium sized vocabulary system. But in the present work, suitable sub-word level acoustical models for small, medium, and large vocabulary systems are proposed. Further, context-dependent tri-phone acoustical models are used to improve the recognition accuracy for the large vocabulary system.

The literature on speech recognition work related to the Indian languages are summarized and tabulated as shown in Table 1. The researchers have used different sub-word acoustical models to improve the system performance. A good word recognition accuracy (above 80%) has been reported in table 1 because of the small vocabulary size considered. Researchers have mainly explored on syllable and phone based acoustical models for different Indian languages such as *Hindi, Tamil, Telugu, Kannada*, *Assamese*, and *Bengali*. The majority of the Indian languages are syllable-timed and researchers have tried exploring this feature. The other challenge is to develop application specific speech database in a particular language as there are no standard database available. In the present work, three different set of speech corpus with vocabulary size of 10, 110 and 1498 words are developed. The small vocabulary consists of utterances of ten *Kannada* digits, the medium sized vocabulary is taken from *Kannada* short story and the large vocabulary is chosen from the *Kannada* broadcast news corpus. The entire speech database is developed by recording the speech utterances from 3 male and 3 female speakers. The objective of this study is to propose suitable sub-word level acoustical models for the different vocabulary based isolated word recognition systems. Also context dependent tri-phone based speech recognition system has been implemented and its performance against context independent mono-phone based systems has been analyzed. So, the novelty of this work mainly includes building of *Kannada* speech database, developing the phone and syllable based dictionaries, and performance evaluation of word recognition system for different vocabulary size.

Table 1
Summary of literature on IWR for Indian languages

Language/ Authors	Vocabulary	Number of Speakers	Acoustical Models Used	Accuracy (%)
Hindi (Aggarwal & Dave, 2011)	400 words	1	Gaussian mixture, Mono-phone HMM	80
Hindi (Saini, Kaur, & Dua, 2013)	113 words	9	Mono-phone HMM	90
Tamil (Radha, 2012)	50 words	10	Mono-phone HMM	88
Telugu (Sunitha, & Kalyani, 2012)	40 words	1	Syllable based acoustic modelling	80
Assamese (Bharali, & Kalita, 2015)	10 words	15	Word level HMM with different hidden states.	80
Kannada (Hegde, Achary, & Shetty, 2012)	10 words	5	SVM classifier	79
Kannada (Thalengala, & Shama, 2016)	1500 words	6	Syllable and phone based HMM	40 and 61

Kannada Language

India is basically multi-lingual country with more than 1500 spoken languages and about 150 languages have sizable speaking population. There are twenty-two official languages recognized by the Indian government and *Kannada* language is one of them. *Kannada* is one of the Dravidian languages of India (Krishnamurti, 2003) and is the state language of Karnataka state with the history of more than 2000 years. As per the recent survey there are about sixty million *Kannada* speaking peoples present inside and outside of Karnataka state. *Kannada* is a syllable-timed language having 52 basic alphabets (called "*Akshara Maala*") which are basically evolved from the "*Kadamba*" script. The alphabets of *Kannada* are grouped into three categories which are named as "*Swaragalu*" (vowels), "*Vyanjanagalu*" (consonants), and "*Yogavahakagalu*" (nasal like consonants) (Hegde, Achary, & Shetty, 2015; Krishnamurti, 2003). The organization of these "*Aksharas*" is well structured and arranged as per the place of articulations. The meaningful sequence of syllable-like alphabets gives rise to words in *Kannada* language. In *Kannada* language, the alpha-syllabary ('*Aksharas*') units are very stable and have unique pronunciations which are independent of their occurrence in a word or sentence. Researchers have tried to explore this characteristics in speech recognition systems (Hegde, Achary, & Shetty, 2012; Hegde, Achary, & Shetty, 2015).

Isolated Word Recognition (IWR) System

Any speech recognition system building has got two major stages viz. training stage and testing stage. The training phase consists of building acoustical models (HMM based sub-word models) from the input training speech data set and is as shown in Figure 1a. Figure 1b shows testing phase, where recognition of input testing data set is carried out by using the sub-word models generated in training stage.

The front-end of the speech recognition system comprises speech pre-processing and cepstral analysis operations which are common to both the training and testing phases. The initial pre-processing step involves pre-emphasis and speech framing operations. It is assumed that the speech is recorded in a controlled environment with minimum noise

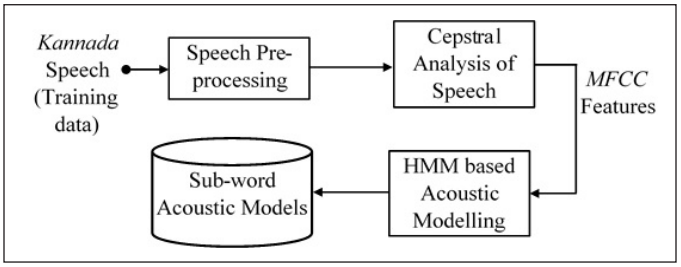


Figure 1a. Acoustic model building steps: The training phase

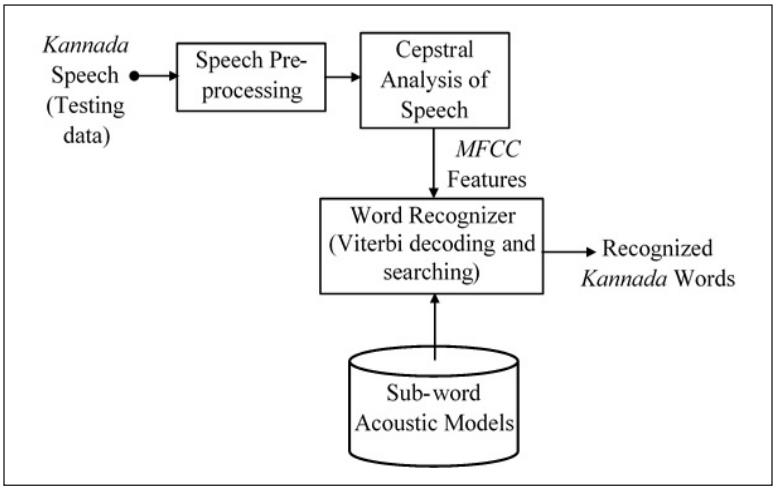


Figure 1b. Speech recognition steps: The testing phase

interference and uniform recording conditions. Speech signal has a characteristic that high frequency components have comparatively less energy than that of the low frequency components. This imbalance can be minimized by passing the speech signal through a pre-emphasis filter (Deller, Proakis, & Hansen, 1993). The spectrally flattened speech signal is segmented into overlapping time frames to obtain speech "frames" or "windows" and all the further processing is carried out on these speech frames. Speech is non-stationary in nature and therefore windowing is required to minimize the effect of non-stationary characteristics. The larger the speech frame length the better is the spectral resolution whereas the smaller the speech frame length the better is the time resolution. Typically window length of 1 or 2 pitch periods are considered in the speech analysis (Deller, Proakis, & Hansen, 1993); (Rabiner, Juang, & Yegnanarayana, 2012). In this work, hamming window of 20msec time duration and 10msec overlapping between the adjacent frames are considered.

The next stage in the front end of speech recognition system is speech parameterization. The performance of speech recognition system is very much dependent on the speech parameters considered in the study. Effectiveness of the speech parameters or features depends on how best it can represent acoustic models. Some of the speech features used in speech recognition systems are linear predictive coefficients (LPC), linear prediction cepstral coefficients (LPCC), perceptual linear prediction (PLP), and Mel frequency cepstral coefficients (MFCC). Among the others, MFCC features have shown better performance in speech recognition systems. These MFCC parameters have characteristics that it mimics the human ear perceptual system (Davis & Mermelstein, 1980). The response of human ear for different frequency bands is not linear but it can be well described by Mel-scale filter bank. In this study MFCC features together with its first and second order differences are considered.

Hidden Markov model (HMM) has been very commonly used in speech recognition task due to its ability to effectively model the time-series events. The back-end of the speech recognition system consists of building the HMM based acoustical sub-word units and applying pattern matching techniques to recognize the input utterances. The acoustical sub-word models considered in this study are words, syllables, and phone units of *Kannada* language. These acoustical sub-word units are developed using HMMs. The HMMs are basically the first order Markov process with non-observed states. So, HMM is a double stochastic process with one hidden state sequence which is calculated from the observation sequences. The HMM can be represented mathematically using its parameter set given by,

$$\lambda = (A, B, \pi) \qquad [1]$$

Where 'π' represents initial state distribution vector, '*A*' is state transition probability matrix and '*B*' is the state observation probability distribution vector (Rabiner, 1989; Nilsson, 2005). The most likely estimate of a word is computed by maximizing the a posteriori function according to,

$$\hat{W} = \arg\max_{W \in L} [P(O/W)P(W)] \qquad [2]$$

In the above equation [2], *(P(O|W))* is the probability of having certain observation sequence '*O*' of length '*L*' for a given word model '*W*'. *P(W)* represents the initial probability of a word model. Here observation vector is same as the feature vector and word model '*W*' consists of sequence of sub-word units. The summary of speech recognition process can be given by the following two steps.

Build individual HMM with same number of states for all the sub-word units. This is accomplished by using Baum-Welch algorithm and is known as HMM training.

1. Use the Viterbi decoding technique to find the best matching word. Decoding process uses the HMMs generated in the above step 1 to find the best state sequence for the given test sample. This is known as recognition or testing phase.
2. Finally, the performance of the system is evaluated based on the word recognition accuracy or using the word error rate (WER).

IWR System Implementation Using HTK

The isolated word recognition (IWR) system for *Kannada* words has been implemented using Hidden Markov Toolkit (Young et al., 2006) (HTK version 3.4.1) in the Linux platform. The system implementation involves mainly data preparation, data coding, acoustic modelling and evaluating the performance of the system. These steps are clearly presented in Figure 2.

In the data preparation stage, speech signal was acquired and then pre-processed. In addition to this, required vocabulary was defined and corresponding lexicon dictionary had been developed. In the speech analysis stage (data coding stage), speech signals were processed to obtain the sequence of feature vectors. The specifications such as window size, window type used, and choice of feature vector ere set during this stage. Next stage is the training phase, where speech features used to generate HMMs for each sub-word unit. Here Baum-Welch re-estimation algorithm was applied to obtain the acoustical sub-word models. This step is realized mainly by using the HTK commands "HCOMPV", "HREST", and "HEREST" as shown in Figure 2. Next, pattern matching of test samples with the stored acoustic models were carried out to obtain the recognized words. This is accomplished by using "HVITE" command in HTK. In the last stage, number of words classified correctly were computed and system recognition accuracy was analyzed.

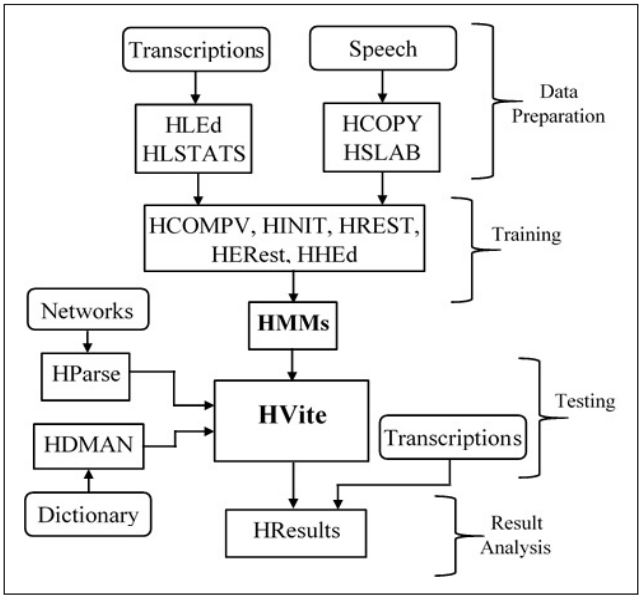


Figure 2. Steps in speech recognition system implementation using HTK

Speech Corpus

The details of speech databases developed for this study has been tabulated in table 2. There are three different speech databases considered in this study. A small vocabulary consisting of utterances of *Kannada* digits, a medium sized vocabulary developed by the word utterances from short *Kannada* story, and a large vocabulary obtained by recording the *Kannada* regional news corpus utterances. For instance, the number of training samples of 540 in Table 2 corresponds to the 10 words uttered by six different speakers nine times

each (10 words * 6 speakers* 9 times = 540 recordings). Similarly, the 180 test samples corresponds to the same 10 words uttered by the same six speakers three times. However the six speakers are different for different databases.

Table 2
Details of speech database developed

Speech Data Considered	Vocabulary Size (Number of Words)	Number of Speakers	Number of Training Samples	Number of Testing Samples
Kannada Digits (0 to 9)	10	6	540	180
Kannada Short Story	110	6	1980	660
Kannada Broadcast News	1498	6	5460	1820

These speech recordings have been carried out in a controlled (minimum noise interference) and uniform (good recording consistency) environment. A linear PCM recorder (Olympus LS-3) has been used in speech recordings and stored in wave format. Sampling frequency of 16 KHz and 16-bit PCM format has been followed while acquiring and digitizing the speech data.

Mono-phone Acoustical Models

The acoustical property and nature of sub-word units are highly influenced by the presence of adjacent sound units. This is mainly because of the co-articulation effect of the human speech production system. Two types of acoustical models can be defined based on the influence of adjacent sound units are context dependent tri-phone models and context independent mono-phone models. In the present study, three different sub-word mono-phone models and a tri-phone model based systems have been implemented. The tri-phone models are basically extended version of mono-phone models. The three mono-phone acoustical models considered in this study are word level model, syllable model, and phone model.

Developing any mono-phone model starts with building appropriate pronunciation dictionary. Both syllable and phone based dictionaries for the *Kannada* words have been developed. The phone set of *Kannada* language considered in this work are given in Table 3 (Shridhara, Banahatti, Narthan, Karjigi, & Kumaraswamy, 2013). There are 50 phones and are labelled with one or more English letters as shown in Table 3. The number of syllables (alpha-syllabary units) needed to represent chosen vocabulary size of 1498 words are found to be 564 syllables.

Syllables are obtained from the meaningful combination of phones and these syllables form alphabets of the *Kannada* language. Examples of how *Kannada* words are represented in pronunciation dictionaries (both phone and syllable based) are shown in Table 4. It

Table 3
Phones in Kannada language

Label	*Kannada* phone	Label	*Kannada* phone	Lable	*Kannada* phone
a	ಅ	g	ಗ್	p	ಪ್
A	ಆ	gh	ಘ್	ph	ಫ್
i	ಇ	ng	ಙ್	b	ಬ್
I	ಈ	c	ಚ್	bh	ಭ್
u	ಉ	ch	ಛ್	m	ಮ್
U	ಊ	j	ಜ್	y	ಯ್
ru	ಋ	jh	ಝ್	r	ರ್
rU	ೠ	ny	ಞ್	l	ಲ್
e	ಎ	T	ಟ್	v	ವ್
E	ಏ	Th	ಠ್	sh	ಶ್
Ai	ಐ	D	ಡ್	ss	ಷ್
o	ಒ	Dh	ಢ್	s	ಸ್
O	ಓ	N	ಣ್	h	ಹ್
Au	ಔ	t	ತ್	L	ಳ್
aom	ಅಂ	th	ಥ್	---	---
aH	ಅಃ	d	ದ್	---	---
k	ಕ್	dh	ಧ್	---	---
kh	ಖ್	n	ನ್	---	---

can be seen from the Table 4, that each *Kannada* word is represented as a sequence of *Kannada* syllables and sequence of *Kannada* phones. Clearly maximum number of phones required to represent the *Kannada* words is limited to 50, whereas number of syllables increases with vocabulary size up to a certain extent. The third acoustical model that has been considered in this study is word itself. This is used only for small vocabulary of 10 words (*Kannada* digit utterances). Word model has significantly poor performance and therefore not considered for larger databases.

Signal processing operations are performed on the speech waveform to obtain the cepstral parameters. It has been proven that Mel frequency cepstral coefficients (MFCC) are good parameters for representing acoustic units of speech (Davis & Mermelstein, 1980). Here speech waveforms are processed to obtain 12-length MFCC feature vectors. Along with the 12-MFCC parameters, signal energy component also included for the better representation. The human speech perception is greatly influenced by the spectral transitions present in the speech signal. This information can be quantified by considering the time difference (first order time derivatives) and acceleration coefficients (second order

time derivatives). So by combining all the parameters (12-MFCC, 1-log energy value, 13-delta coefficients, and 13-delta-delta coefficients) together the length of the feature vector becomes 39.

Table 4
Examples of phone and syllable sequences for Kannada words

Kannada word	English meaning	Sequences of syllables	Sequence of phones
ಸಹಕಾರ (sahakara)	Help	Sa ha kA ra	s a h a k A r a
ಮೂಲಕ (moolaka)	Through	mU la ka	m U l a k a
ಸಾವಿರ (savira)	Thousand	sA vi ra	s A v i r a
ಅವರು (avaru)	They	A va ru	a v a r u
ನೂತನ (noothana)	New	nU ta na	n U t a n a

Speech windows (Hamming window) of length 20 msec with 10 msec overlapping between the adjacent frames are considered in this study. The speech coding specifications used in HTK implementation (using "HCOPY" command) is tabulated as in Table 5. The speech parameterization is done for entire speech database by considering the both the training and testing speech samples. The feature vectors belonging to training set are used to develop the HMMs, whereas the one belonging to testing set are used to assess the models.

Table 5
Speech coding specifications

Parameters	Specifications
Sampling rate	16000 Hz
File format	16-bit PCM, mono, wave file
Window length	20 msec
Window used	Hamming
Window overlapping	50%
Pre-emphasis coefficient	0.97
Features derived	MFCC, ΔMFCC, and ΔΔMFCC
Feature vector length	39
Number of filter-banks	26
Number of MFCC coefficients	12

Context-independent mono-phone models have been built from the training feature set using the HMMs. This is implemented mainly by using "HCOMPV", "HREST", and "HEREST" commands in HTK tool as shown in figure 2 (Young et al., 2006). All the three acoustical models namely word-level, syllable-level and phone-level are considered in the design of mono-phone models. All the sub-word units are built by using 5-state HMM with first and last states are being non-emitting states. The HMM states are considered to be Gaussian distributed with first and last states representing the word boundaries (silence states). Initially the flat start prototype mono-phone models for each sub-word unit in the dictionary are defined. It is assumed that all the states in HMM are of Gaussian type with zero mean and unit variance. Features from the training data set are used to estimate and re-estimate the HMM based sub-word models. This is accomplished by the "HEREST" function, which adapts Baum-Welch re-estimation algorithm to obtain the trained HMMs. There are 50 phone models and about 500 syllable models have been developed in this study. The Baum-Welch re-estimation is a forward-backward algorithm which includes the pruning limit in its summation to reduce the amount of computations. Tight pruning thresholds can be kept for most of the training data sets, however, some training data set may show poor acoustic matching and as a result wider pruning threshold is needed. This issue is addressed by using an auto-increment pruning threshold. In the present study, pruning threshold of value of 250 has been considered, and if re-estimation fails, the threshold is incremented by a value 150. This is repeated until either the data is successfully processed or the pruning limit value of 1000 is exceeded.

After the training stage, recognition accuracy of the system evaluated using the testing data set. The Viterbi decoding technique ("HVITE" command in HTK) has been applied to calculate the best matching word from the dictionary for every testing sample. Based on the pattern matching results, recognition accuracy is obtained (using "HRESULTS" function).

Tri-phone Acoustical Models

The acoustical characteristics of a phone is highly influenced by the neighboring phones in any language. For instance, the utterance of the phone */a/* in the word "*acoustic*" sounds different for the same phone in the word "*automatic*". This is basically due to the co-articulation effects of the human speech production system. So, acoustical properties of phones are dependent on the occurrences of adjacent phones and hence context-dependent phones are to be defined. In tri-phone modeling, acoustical models for each phone is obtained by considering the effect of adjacent (immediate left and immediate right) phones. The tri-phone units are obtained from the phone sequences by concatenating the left and right phones. For a phone sequence "*th ih s*" (for the word "*this*"), the resultant tri-phone sequence would be "*th+ih th-ih+s ih-s*". The "+" and "-" symbols are used representing the right and left contexts of a phone respectively. Observe that the word boundaries gives

rise to bi-phone units. The conversion from mono-phone into tri-phone has got two steps. First, mono-phone transcriptions are cloned and converted into tri-phone transcriptions and tri-phone models are re-estimated. In the second step, the similar acoustic states of tri-phones are tied so that all state distributions are robustly estimated (Young et al., 2006). These implementation steps in HTK tool are summarized in Figure 3.

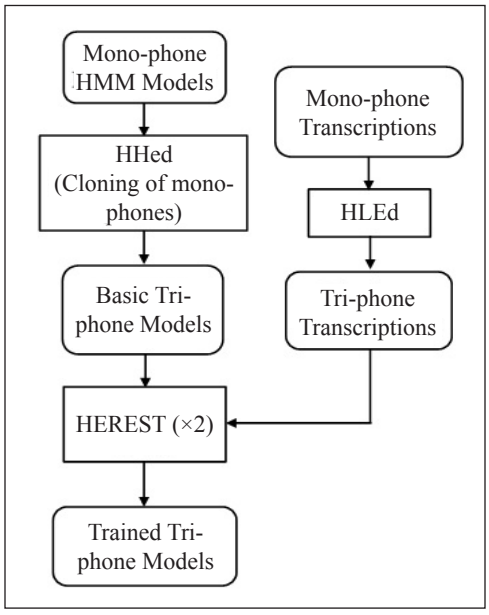


Figure 3. Mono-phones to Tri-phone conversion steps

In the previous step, the tri-phones generated such that the phone set share the same transition matrix. But the recognizer performance mainly depends on the accuracy of state output distributions that represent the input speech acoustics. So, the final stage in the tri-phone model building is to tie the states within tri-phones so that it is possible to make robust parameter estimation. This is achieved by 'HHED' command which clusters the states within tri-phones using decision tree method. The decision based technique requires questions to be set up regarding the left and right contexts of each tri-phones. Each question describes one set or class of contexts. For example a question named "left-stop-class" may define all possible "stops" to the left of any phone. The questions are defined such a way that they must be able to address the wide general class of sound units such as vowels, consonants, nasals and so on. Further, the questions can also include certain linguistic or phonetic classifications as appropriate. In the present tri-phone implementation the general class of phones are considered. But there is further scope to tune the tri-phones by exploring on the contextual occurrences of *Kannada* phone units. Once all the state-tying has been carried-out, the more refined tri-phone models are again re-estimated (using "HEREST" command) to obtain the final tri-phone models.

Finally, the word recognition was carried out by using Viterbi decoding algorithm (using "HVITE" command which is the same as given in Figure 2) on the test data set.

RESULTS AND DISCUSSION

The performance of any speech recognition system in general is evaluated by using word recognition accuracy. The word recognition accuracy is defined as follows.

$$\text{Recognition accuracy} = \frac{N-S}{N} \times 100\% \qquad [3]$$

Where '*N*' represents the number of words in the test data set and '*S*' is the number of words replaced (known as substitution error). Speech database considered here are obtained from the recordings of six different speakers and these databases have been pooled into three groups for experimentation. Each group has two speaker's data with one male and one female speaker. The well-known "*holdout*" procedure has been adapted in this analysis where one of the data group is used for testing and the remaining two for training. All the results presented here are the average of results for three data groups. The overall recognition result obtained in this study is summarized in Table 6. The performance analysis of the IWR system has been done based on the vocabulary size and on the acoustic model adapted. In this paper IWR system for small, medium, and large vocabulary are considered.

Table 6
Results of word recognition accuracy

Vocabulary Size	Acoustical Model Used	Number of HMMs	Accuracy (%)
10	word	10	97.22
10	Syllable	18	98.88
10	Phone	21	98.33
50	Syllable	93	96.33
50	phone	35	95.33
110	Syllable	138	95.15
110	Phone	37	94.24
921	Syllable	441	56.52
921	Phone	49	72.59
1498	Syllable	564	51.62
1498	Phone	49	67.82
921	Tri-phone	2114	82.57
1498	Tri-phone	3033	70.56

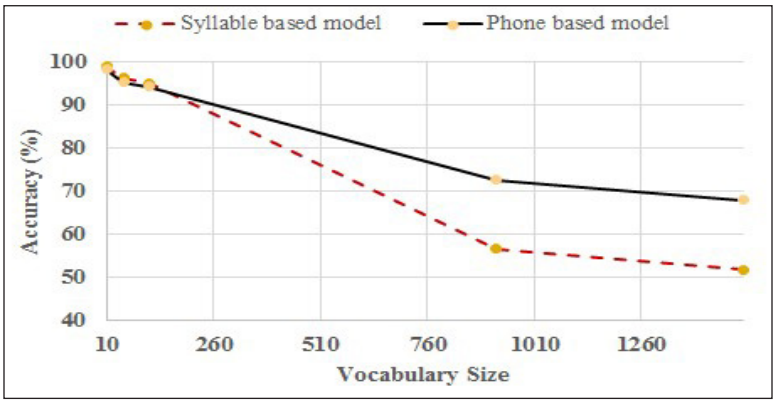


Figure 4. Performance of sub-word acoustical models

The syllable and phone based system performances for different vocabulary size has been plotted in Figure 4. It can be seen that the recognition accuracy of the both phone and syllable based systems decreases with the increase in vocabulary size. But the performance of phone based system is better than that of syllable based systems for larger vocabulary size. It can be observed that phone based system gives about 16% better results than the syllable based system for the vocabulary of 1498 words. The number of syllable models (number of HMMs) required increases as number of words increase, whereas number of phone models are limited to 50 for *Kannada* language. It is observed from the table 6 that forty nine phones are used and one phone left unused for the chosen vocabulary. Also when the acoustical models become larger in number, training that each model undergoes becomes lesser. Therefore we can see that the performance of syllable based system becomes poor with the increase in vocabulary size. Based on these findings, we propose the syllable models for small and medium sized vocabulary systems, whereas phone models for larger vocabulary systems. So the choice of acoustic model becomes application specific.

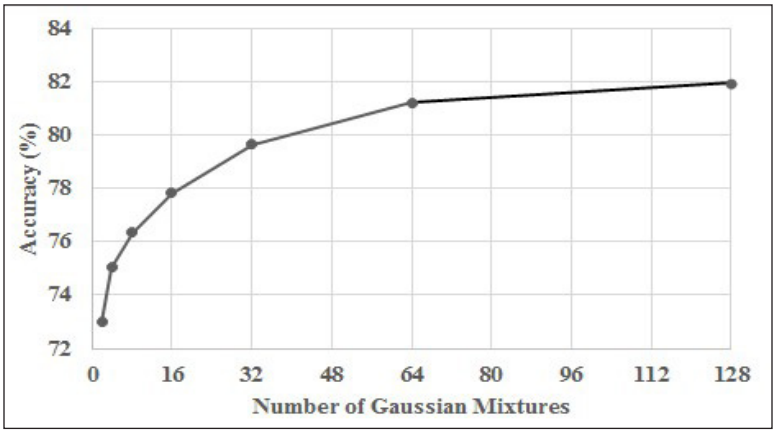


Figure 5. Performance of mono-phone models against number of Gaussian mixtures.

The individual states of HMM are assumed to be Gaussian distributed. For the better representation of sound unit, each HMM state can be made multivariate Gaussian density distributed function. But for the efficient modelling using Gaussian mixtures, a large training database is required. In this study, mono-phone models (phone based models) with Gaussian mixture models (GMM) have been implemented for vocabulary size of 1498 words. The results of recognition accuracy against the number of Gaussian mixtures is plotted in Figure 5. So the mono-phone system performance increases with the number of Gaussian mixtures up to a certain extent. But GMM implementation has a disadvantage that the training phase computational complexity is high and also require large training data set.

Now, all mono-phone models considered (word, syllable, and phone) in this work are context independent acoustic models. This means that each sub-word unit in a mono-phone system is built individually without considering its place of occurrence (context) in a word utterance. But the acoustical property of a sub-word unit in speech is greatly affected by the presence of adjacent sub-word units. In-order to address the phone-to-phone acoustical variations present in the utterance of a word, context dependent tri-phone models have been developed. The results of these tri-phone models are presented in the last two rows of Table 6. The results of context dependent tri-phone models show significant improvement over the mono-phone models. It can also be noted that the performance of GMM based mono-phone reaches near to the tri-phone performance with a mixture size greater than 64. Hence we propose to choose tri-phone models for large vocabulary *Kannada* word recognition systems.

CONCLUSION

Mono-phone and tri-phone based isolated word recognition systems for *Kannada* words have been realized in this study. The implementation of speech recognition system carried out successfully using HTK tool box. Context independent mono-phone systems with three different acoustical models namely word-level, syllable-level, and phone-level have been implemented for different vocabulary size. The implementation involves developing syllable and mono-phone dictionaries for *Kannada* words. Also context independent mono-phone models are extended towards context dependent tri-phone models. Recognition accuracies of these various implementations have been analyzed. The best word recognition accuracies of 67.82% and 70.56% are found respectively for mono-phone and tri-phone based systems on *Kannada* broadcast news database having vocabulary size of 1498 words.

It can be therefore proposed that syllable based acoustical models are suitable for small and medium size vocabulary systems. For applications requiring larger vocabulary, phone based models or context dependent tri-phone models can be chosen. System performance may be improved further by increasing the training data set. Also, there is a further scope for tuning the tri-phone models by incorporation language specific classifications in the questions tree.

REFERENCES

Aggarwal, R. K., & Dave, M. (2011). Using Gaussian mixtures for Hindi speech recognition system. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 4*(4), 157-170.

Ananthakrishna, T., Maithri, M., & Shama, K. (2015, December). *Kannada* word recognition system using HTK. In *2015 Annual India Conference (INDICON)* (pp. 1-5). New Delhi, India.

Bharali, S. S., & Kalita, S. K. (2015). A comparative study of different features for isolated spoken word recognition using HMM with reference to Assamese language. *International Journal of Speech Technology, 18*(4), 673-684.

Bhaskar P. V., Rao R. M. S. & Gopi A. (2012). HTK Based Telugu Speech Recognition.*-International Journal of Advanced Research in Computer Science and Software Engineering, 2*(12), 307-314.

Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 28*(4), 357-366.

Deller J. R., Proakis J. G. & Hansen J. H. L., (1993). *Discrete Time Processing of Speech Signals*. New York: Macmillan Publishing Company.

Dua, M., Aggarwal, R. K., Kadyan, V., & Dua, S. (2012). Punjabi automatic speech recognition using HTK. *IJCSI International Journal of Computer Science Issues, 9*(4), 1694-0814.

Gawali, B. W., Gaikwad, S., Yannawar, P., & Mehrotra, S. C. (2011). Marathi isolated word recognition system using mfcc and dtw features. *ACEEE International Journal on Information Technology, 1*(01), 21-24.

Hegde, S., Achary, K. K., & Shetty, S. (2012). Isolated word recognition for *Kannada* language using support vector machine. In *Wireless Networks and Computational Intelligence* (pp. 262-269). Berlin, Heidelberg: Springer.

Hegde, S., Achary, K. K., & Shetty, S. (2015). Statistical analysis of features and classification of alphasyllabary sounds in *Kannada* language. *International Journal of Speech Technology, 18*(1), 65-75.

Krishnamurti, B. (2003). *The Dravidian Languages*. Cambridge: Cambridge University Press.

Kumar, K., Aggarwal, R. K., & Jain, A. (2012). A Hindi speech recognition system for connected words using HTK. *International Journal of Computational Systems Engineering, 1*(1), 25-32.

Kumar, M., Rajput, N., & Verma, A. (2004). A large-vocabulary continuous speech recognition system for Hindi. *IBM Journal of Research and Development*, *48*(5.6), 703-715.

Lakshmi, A., & Murthy, H. A. (2006). A syllable based continuous speech recognizer for Tamil. In *Ninth International Conference on Spoken Language Processing* (pp. 1878-1881). Pittsburgh, Pennsylvania.

Nilsson, M. (2005). *First Order Hidden Markov Model: Theory and implementation issues*. Technical Report 2005:02. Blekinge Institute of Technology.

Panda, S. P., & Nayak, A. K. (2016). Automatic speech segmentation in syllable centric speech recognition system. *International Journal of Speech Technology, 19*(1), 9-18.

Punitha, P., & Hemakumar, G. (2014, March). Speaker dependent continuous *Kannada* speech recognition using HMM. In *2014 International Conference on Intelligent Computing Applications (ICICA),* (pp. 402-405). Coimbatore, India.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77*(2), 257-286.

Rabiner, L. R., Juang B. H., & Yegnanarayana B. (2012). *Fundamentals of speech recognition*. Englewood Cliffs: PTR Prentice Hall.

Radha, V. (2012). Speaker independent isolated speech recognition system for Tamil language using HMM. *Procedia Engineering, 30*, 1097-1102.

Saini, P., Kaur, P., & Dua, M. (2013). Hindi automatic speech recognition using HTK. *International Journal of Engineering Trends and Technology (IJETT), 4*(6), 2223-29.

Shridhara, M. V., Banahatti, B. K., Narthan, L., Karjigi, V., & Kumaraswamy, R. (2013, November). Development of *Kannada* speech corpus for prosodically guided phonetic search engine. In *Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE), 2013 International Conference* (pp. 1-6). Gurgaon, India.

Sunitha, K. V. N., & Kalyani, N. (2012). Isolated Word Recognition using Morph Knowledge for Telugu Language. *International Journal of Computer Applications, 38*(12), 47-54.

Thalengala, A., & Shama, K. (2016). Study of sub-word acoustical models for *Kannada* isolated word recognition system. *International Journal of Speech Technology, 19*(4), 817-826.

Thangarajan, R., Natarajan, A. M., & Selvam, M. (2009). Syllable modeling in continuous speech recognition for Tamil language. *International Journal of Speech Technology, 12*(1), 47-57.

Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X & Valtchev, V. (2006). *The HTK book*. Cambridge, UK: Cambridge University Press.

Pertanika J. Sci. & Technol. 26 (4): 1867 - 1880 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Interference Coordination for Multiple Resource Sharing in D2D Communication Underlaying LTE-A Network

Sharma, S.* and Singh, B.

Department of Electronics and Communication Engineering, National Institute of Technology, Kurukshetra-136119, India

ABSTRACT

Data traffic has been increasing at an exponential rate causing extremely heavy demand on cellular networks. Device-to-Device (D2D) communication is seen as a potential solution for data traffic offloading and enhanced performance of cellular networks. It improves the spectrum and energy efficiency of the network. But co-channel interference is the major concern while performing spectrum sharing in D2D communication. Most of the existing works have proposed power control schemes for non-overlapping spectrum allocation among D2D pairs. This paper focuses on reducing the co-channel interference by dynamic power control while allocating multiple resources to single D2D user. D2D system performance is formulated as an overlapping coalition game coupled with interference based transmit power distribution among the resource blocks assigned to a single user. Simulation results show that the proposed scheme outperforms the other existing techniques in terms of D2D throughput and total transmit power

Keywords: Device-to-Device communication, overlapping coalition game, power control, resource allocation

INTRODUCTION

The cellular data traffic has been increasing at an exponential rate as number of smart phone users are growing rapidly. The global mobile data traffic is expected to grow by more than200 times over the period from 2010 through 2020 and by nearly 20,000 times from 2010 to 2030. In order to accommodate the ever growing wireless users, provide better quality of service for applications like online gaming, social networking and video streaming. there is an imperative need to explore new radio spectrum enabling greater bandwidth and efficient radio access technologies. Basically there are two

ARTICLE INFO

Article history:
Received: 16 January 2018
Accepted: 28 May 2018
Published: 24 October 2018

E-mail addresses:
sandeepika31@gmail.com (Sharma, S.)
brahmjit@nitkkr.ac.in (Singh, B.)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

approaches to increase the throughput and spectral efficiency of a cellular network, either to shift to a new broader spectrum or to efficiently utilize the available spectrum resources. To adopt the former approach researchers are exploring the millimeter wave (mmWave) band for next generation cellular communication. But at high operating frequencies omni-directional path loss is high as compared to conventional sub 3 GHz band hence making long distance communication difficult at millimeter frequencies. Also mmWave signals are very susceptible to shadowing, resulting in outages and reduced channel quality. Therefore to improve the utilization of available radio resources is an alternative approach to it which utilizes technologies like small cell deployment, heterogeneous network, redesigning backhaul links and spectrum sharing. To perform spectrum sharing, either transmission parameters can be adapted according to the characteristics of the environment in which it operates as done in cognitive radio or traffic may be offloaded to underutilized licensed and unlicensed spectrum as done in device-to-device (D2D) communication underlaying cellular networks. D2D communication is a potential solution to enhance the reuse of radio resources facilitating high throughput, improved spectrum and energy efficiency of mobile multimedia in 5G networks, and extended network coverage. But as implementing D2D communication requires introduction of new functionalities both at user device and base station, it significantly complicates the cellular network design. Hence D2D communication along with cellular networks has attracted significant attention and its feasibility in context of LTE-A is being investigated through academic and industrial research.

First implementation for D2D communication improving spectral efficiency as compared to the 802.11 architecture with RTS/CTS was proposed by Qualcomm (Wu et al., 2013). This concept was adopted for U-D2D communication to IMT-A and LTE-A cellular network, providing high speed, low cost and seamless connectivity for rich multimedia applications (Doppler et al., 2009; Feng et al., 2013; Janis et al., 2009). But various design issues like D2D mode selection, channel quality estimation with minimum signaling overhead, D2D bearer establishment, best selection of UL/DL resources for sharing, scheduling interaction time between network and D2D pair and optimal power control were encountered in realizing it (Fodor et al., 2012 ; Wang & Tang, 2016). To overcome these challenges various mode selection, resource allocation and interference management techniques have been proposed.

Resource sharing (uplink, downlink or both) with cellular users introduces problem of co-channel interference which limits overall performance of the cellular network and maximum number of D2D links that can be supported along with cellular links. Co-channel interference may be co-tier i.e. among the D2D pairs sharing same RBs and cross-tier which arises between cellular and D2D pair sharing same RB. In literature, several works have been proposed to alleviate this problem through optimal resource allocation, joint strategies either for mode selection and resource allocation or for resource and power

allocation (Sobhi-Givi, Khazali, Kalbkhani, Shayesteh, & Solouk, 2017). Location based mode selection and resource allocation were performed in Rodziewicz (2015). Advantage of such scheme is that it does not require any channel state information. Power control and resource allocation is optimized jointly for a predefined QoS, based on the type of application like video streaming, and file sharing in (Sun, Zhang, Liang, Zhang, & Chen, 2016). In Ma, Liu and Jiang, (2016) it was performed by calculating interference limited area (ILA) for each cellular user and D2D pair separately. Various game theoretic and matching techniques like Stackelberg game, reverse iterative combinatorial auction model, one-to-one matching, bipartite matching etc. have also been utilized to model and analyze the behavior of D2D users effectively (Li et.al., 2014; Song et al., 2014 ; Xu et.al., 2012; Xu et.al., 2014). Social context aware D2D architecture has been proposed by researchers (Ciou et al., Li et al., 2014; 2015; Semiari et al., 2015). Influence of social ties and energy consumption are current concerns arising in D2D networking scenario (Datsika, Antonopoulos, Zorba, & Erikoukis, 2016). D2D is also seen as a promising solution for Fog Radio Access Network (F-RAN). In F-RANs popular contents are stored in the mobile device and it is shared among other devices. Here social relationship strength maximizes the system utility and effectively alleviates the burden of front-haul (Zhang, Sun, Mo, Zhang, & Bu, 2016). To realize such group communication authors have suggested D2D communication in multi-cell architecture by using fractional frequency reuse (Jiang, Wang, Sun, Liu, & Wang, 2016).

Relaxing this constraint of allocating single RB to a D2D pair some work has focused on multiple resource sharing improving system sum rate up to 64% relative to one-to-one matching approaches adopted earlier. It also increases the number of successful D2D transmission by 60% (Zhao, Liu, Chai, Chen, & Elkashlan, 2017). Generally D2D User Equipment (DUE) transmit power used over all RBs allocated to a single D2D pair is obtained through uniform power distribution method i.e. maximum DUE transmit power is equally distributed among all the RBs allocated to it (Wang et al., 2011; Xiao et al., 2015; Xu et al., 2016; Xia et al., 2015).However, it is analyzed in literature that co-tier interference is a major reason to lower the potential gain of U-D2D. Therefore an attempt to reduce co-tier interference need to be made in multiple RB allocation mode through channel based power control. This method of power control has been used earlier to improve the performance of conventional OFDM systems (Cover & Thomas, 2012; Kim et al., 2005). They have shown significant enhancement in system capacity by utilizing iterative water-filling scheme for channel based power allocation

In this paper we investigate the problem of interference management in multiple RB allocation mode by proposing an overlapping coalition formation game framework coupled with dynamic control. Earlier works reported in literature in U-D2D as observed by us have not considered interference based power control scheme in multiple resource allocation

scenario. Our contributions are summarized as follows. Multiple resource allocation to a single D2D pair in U-D2D scenario is realized by formulating an overlapping coalition framework. D2D throughput is enhanced by optimizing the resource allocation and dynamically controlling the transmit power a D2D transmitter based on the interference experienced over each channel or RB allocated to it. The fractions of total transmit power (considered as resource residing with D2D pairs) contributed by a D2D pair to each coalition so formed is decided through modified Water-Filling approach. This results in higher system sum rate and D2D transmission ratio.

System Model

A single cell scenario is considered with BS situated at the centre, cellular users and direct D2D pairs distributed uniformly in the cell coverage area as shown in Fig.1. Each user device is equipped with a single antenna therefore a D2D pair can operate in half duplex mode only. It is assumed that at any instant, there are N cellular user equipments (CUEs) and K D2D pairs in the cell. Set of CUEs and D2D pairs can be represented as $C = \{C_1, C_2, \dots C_N\}$ and $D = \{D_1, D_2, \dots D_K\}$ respectively. Though whole uplink bandwidth is available for reuse, we assume that the available RBs are equal to those used by cellular users in the cell. We denote set of RBs as $B = \{B_1, B_2, \dots B_K\}$. Each CUE from set C is assigned an RB from set B following similar indexing i.e. orthogonal resource allocation. As a CUE can share same RB with multiple D2D pairs, this causes interference at BS due to D2D transmission and at DUE receivers due to cellular uplink transmission. Element x_{C_i,D_i} is used to indicate sharing of B_i among cellular user and D2D pairs. If C_i shares a RB with D_j then $x_{C_i,D_i} = 1$. Element x_{D_i,D_j} is used to indicate mutual sharing of RBs among D2D pairs. If $x_{D_i,D_j} = 1$ then D_i is sharing same RB with D_j. Also one D2D pair may occupy multiple RBs such that SINR at all DUEs and CUEs sharing those RB is maintained above a threshold value.

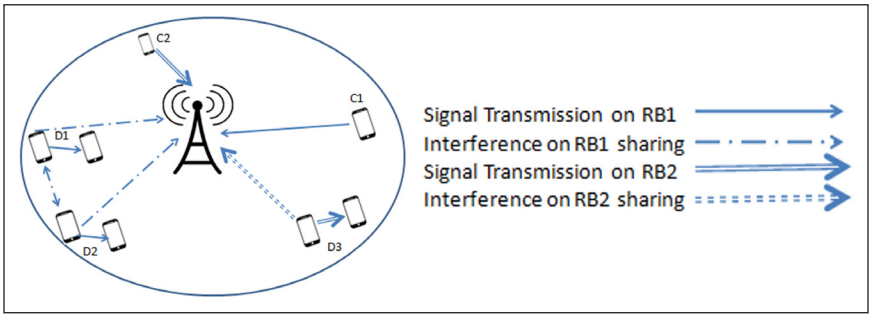


Figure 1. CUEs sharing RBs with D2D pairs causing uplink and mutual interference

Link quality is determined by computing SINR for cellular and D2D transmission links. It is considered as an important parameter for system performance evaluation. Effect of co-channel interference is also accounted for in SINR calculation by

$$SINR_{C_i} = \frac{P_{CUE} G_{C,BS}}{P_{noise} + \sum_{j=1}^{K} x_{C_i,D_j} P_{B_i,D_j} G_{C_i,D_j}} \quad (1)$$

$$SINR_{B_i,D_j} = \frac{P_{B_i,D_j} G_{D_j}}{P_{noise} + P_{CUE} G_{C_i,D_j} + \sum_{i=1}^{K} x_{D_i,D_j} P_{B_i,D_j} G_{D_i,D_j}} \quad (2)$$

Where, P_{CUE} denotes the transmit power of a CUE and P_{B_i,D_j} denotes DUE transmitter power of D_j while using B_i. P_{D_j} is the total transmit power at D_j. $G_{C,BS}$ is the channel gain between C_i and BS . G_{D_j} is the channel gain between DUEs forming D_j, G_{C_i,D_j} is the channel gain between CUE and D2D pair receiver sharing the same RB and G_{D_i,D_j} is the channel gain between transmitter DUE and receiver DUE of D2D pairs sharing the same RB. P_{noise} is noise power.

According to Shannon-Hartley theorem, data rates on j^{th} D2D pair using B_i can be calculated as follows

$$R_D^i = B\left(1 + log_2\, SINR_{B_i,D_j}\right) \quad (3)$$

Where, B is sub channel bandwidth and R_D^i is the D2D data rate while using i^{th} resource block.

Considering all D2D and cellular users, the gain of system corresponds to a total D2D user throughput which can be computed as follows

$$U(R_D) = \sum_{i \in C} \sum_{j \in D} x_{C_i,D_j} R_D^i \quad (4)$$

The value of R_D depends on (i) sharing relation among different D2D pairs and between cellular user and D2D pairs (ii) co-channel interference. By devising optimal resource allocation and power control mechanism, system sum rate can be maximized. Sum rate maximization can be formulated as an optimization problem given below.

$$\max\ U(R_D),$$

$$\text{s.t.} \begin{cases} x_{C_i,D_j} \in \{0,1\}\ \forall i \in \{1,2,\dots N\}, \forall j \in \{1,2,\dots K\} \\ \sum_{i \in B} P_{B_i,D_j} \le P_{D_j}, \forall j \in \{1,2,\dots K\} \\ R_{C_i} \ge R_c^{th}, \forall i \in \{1,2,\dots N\} \\ R_{D_j} \ge R_D^{th}, \forall j \in \{1,2,\dots K\} \\ \sum_{i \in B} x_{C_i,D_j} \le N, \forall j \in \{1,2,\dots K\} \end{cases} \quad (5)$$

R_c^{th} and R_D^{th} are the threshold data rate values for C_i and D_j respectively which guarantees QoS requirements of the system . To limit the interference at each RB and reduce implementation complexity the maximal number of RBs that are allowed to be allocated to a single D2D pair can be restricted based on type of information that is to be shared among the D2D pairs. Here it is assumed to be equal to total number of RBs available. Further, overlapping coalition formation game theory is applied onto the scenario to maximize system sum rate under the given constraints.

Overlapping Coalition Formation for Multiple Resource Allocation

Overlapping coalition formation game theory is used to model practical scenarios in which players may belong to more than one coalition and contribute their resources among those coalitions to improve overall system performance (Wang, Song, Saad, & Han, 2016). The eNB forms cooperative groups termed as coalition based on the signal-to-noise ratio (SINR) feedback from DUEs so as to reduce the co-channel interference between D2D users and cellular users sharing the same RB. The maximum number of coalitions that can be formed is the total number of uplink channels (RBs) available for sharing. Hence K D2D pairs pairs in the given scenario which can form at most 'N' coalition when generally $N < K$. The aim is to derive a stable coalition structure which provides improved system utility value defined in terms of D2D user throughput as given in Eq.(5).

As it is assumed that a D2D pair can be allocated multiple RBs, therefore total DUE transmit power needs to be distributed among the number of RBs allocated to that pair. But as a user may experience different channel gain on different RBs therefore, uniform distributions of total DUE transmit power over each RB may not provide the desired performance. Hence, unlike earlier works which have assumed uniform distribution of DUE transmit power, the proposed approach distributes it in accordance with the interference experienced during that RB. Therefore, the term P_{B_i,D_j} is introduced to denote transmit power used by D_j while transmitting over B_i.

Definition 1. *(Overlapping Coalition Game for Resource Allocation):* The proposed overlapping coalition game $G = (D, R_D, F)$ with a transferable utility (TU), player set $D = \{D_1, D_2, \dots D_K\}$ and coalition structure $F = \{F_1, F_2, \dots F_N\}$, is given by a function given below

$$v = [0,1]^K \rightarrow R_{D+} \tag{6}$$

Such that $\begin{cases} v(0^K) = 0 \\ v(F_i) = \sum_{j \in D} R_D^i \end{cases}$

$v(F_i)$ is the value for every coalition F_i which is the subset of D and a transferable utility with the aim to maximize the system sum rate. This is distributed among the members of F_i depending upon the fraction of power contributed by individual member. Total payoff value of a D2D pair D_j, is the sum throughput achieved by it being the member of F_j^* where $F_j^* \subseteq F$ (F_j^* represents those coalitions which has D_j as its member).

In the proposed game D2D pairs are motivated to join a coalition to increase their own utility value given by $v\left(D_j^{F_j^*}\right)$ and coalition value $v(F_i)$ given in (6). Utility value of each D2D pair is expressed as

$$v\left(D_j^{F_j^*}\right) = \sum_{i \in F_j^*} R_D^i \tag{7}$$

Hence a preference order can be defined for each player and for each RB separately. That preference order can be used to perform split and join operation respectively, which leads to a stable coalition structure offering system sum-rate maximization.

Co-Channel Interfrence Based Power Control

As number of users sharing the same RB increases, co-tier interference experienced by D2D pairs over that RB also increases. Therefore separate power constraints are set for each RB allocated to a single D2D pair. Then P_{B_i,D_j} can be computed using modified water-filling approach as follows

$$P_{B_i,D_j} = \left[\Delta_{D_j} - \left(\alpha\left(\frac{1}{G_{B_i,D_j}}\right)\right)\right]^+ \tag{8}$$

$$\Delta_{D_j} = \frac{1}{g}\left(P_{D_j} + \sum_{G_{B_i,D_j} \in G_j} \alpha\left(\frac{1}{G_{B_i,D_j}}\right)\right) \tag{9}$$

$$G_{B_i,D_j} = \frac{\left|H_{B_i,D_j}\right|^2}{N_i} \tag{10}$$

Here 'α' is the regularization parameter which is used to scale the impact of interference for resource sharing relation among D2D pairs. G_j is the set of RBs allocated to j^{th} D2D pair and 'g' is the cardinality of G_j and $[b^+] = \max(b, 0)$. G_{B_i,D_j} is the SINR value at D_j while using B_i. H_{B_i,D_j} is channel impulse response and N_i is noise due to interference on B_i.

Proposed Algorithm (OCF-PC)

The proposed scheme is a two stage algorithm which comprises initialization and coalition structure update process as explained below

(a) Initialization: When D2D users request for channel assignment, eNodeB prepares a preference list of suitable RBs which can be reused for D2D communication based on the SINR achieved over each channel or RB as a feedback from D2D pairs. Such a preference list is prepared separately for each D2D pair and is valid for one session. In the initialization phase eNodeB allocates the best RB to each D2D pair. Number of coalition formed in this stage is equal to number of RBs available and users sharing a RB are the respective coalition members.

(b) Preference List Update Process: This process aims to update associated RB preference list for each D2D pair so as to maximize the overall utility value i.e. $U(R_D)$. Every coalition is updated through leave and join operations performed for each D2D pair. As overlapping coalition theory is used hence single D2D pair may join more than one coalition. The choice of joining a coalition is motivated by improvement in its own data rate and utility value of that coalition. D2D transmit power is adjusted over each RB after joining and leaving process.

The update process is repeated for each D2D pair until a stable coalition structure is attained.

Table1
Multiple resource allocation and power control (OCF-PC)

Input: 'N' cellular users and available RBs, 'K' D2D pairs, maximum D2D distance
Output: D2D user throughput

Begin
1. Initialization:
• D2D users request RBs from eNodeB.
• eNodeB prepares a preference list of RBs available for reuse based on the SINR feedback from each user. Each D2D pair is allocated the best RB.
• This provides an initial coalition structure.
2. Preference List Update:
For each D2D user D_j
For each F_i if $F_j^* \cap F_i = \phi$
If $v\left(D_j^{F_j^*}\right) > v\left(D_j^{F_j^* \cup F_i}\right)$
and $v(F_i) > v(F_i \cup D_j)$
Then join F_i
Else
Leave F_i

Table1 *(continue)*

eNodeB resets P_{B_i,D_j} by using (8)
End For
End For

3. Compute $U(R_D)$

4. Repeat step 2 and 3 to achieve maximum $U(R_D)$ and a stable coalition structure
End

Properties of Proposed Algorithm.

(a) Stability. In the proposed overlapping coalition game $G = (D, R_D, F)$, if no player $j \in D$ can join or leave any coalition provided the sharing relations of other users, then G has attained a stable coalition structure referred here as Closure-stable (C-stable). In game G, total number of possible coalitions for a given set of D2D pairs and RBs is finite. Hence the sequence of CS-update process will terminate with probability 1 and algorithm converges to $F_{closure}$ (C-stable coalition structure). Suppose if $F_{closure}$ obtained from the given algorithm is not C-stable. Then there exist a D2D pair D_j for has higher preference for F_j^* than F_i. But in our algorithm, such a pair can join F_i with probability 1. This contradicts $F_{closure}$ being the final coalition structure.

(b) Convergence. Starting from an initial coalition structure the proposed algorithm converges to a C- stable coalition structure after limited number of iterations. If F_0 is an initial coalition structure, after an update process in which all users try to maximize their utility, the change in coalition structure can be expressed as $F_0 \rightarrow F_1 \rightarrow F_2 \cdots$.

After m^{th} iteration in an update process, F_{m-1} changes to F_m and for any user $j \in D$, F_j^* changes from $F_{m-1,j}^*$ to $F_{m,j}^*$ satisfying following equation

$$v\left(D_j^{F_{m,j}^*}\right) \geq D_j^{F_{m-1,j}^*} \tag{11}$$

If this change is a result of a D2D pair leaving a coalition then transmit power for every other user of that coalition increases as the co-channel interference reduces by a value G_{B_i,D_j}. Whereas if the change in coalition structure is a result of a D2D user joining a coalition then it can be shown that system utility increases.

(c) Complexity. The complexity of the proposed algorithm depends on the number of leave and join operations in one update process.The exhaustive optimal algorithm for the defined

game has the complexity of $O(N^K)$. Whereas the proposed algorithm has the worst case complexity of $O(KN)$ this is similar to the complexity of scheme utilizing overlapping coalition without power control. Hence it can be said that power distribution function does not add any complexity to the resource allocation algorithm.

Performance Analysis

To evaluate the performance of proposed algorithm, a single cell of radius 500 m is considered with eNB located at its centre. Cellular users and D2D pairs are distributed uniformly in the network area under consideration. Number of cellular users and thus the uplink resources are fixed a prior. Fixed number of available resources presents an upper limit on maximum number of D2D pairs which can share those. This is so because increase in 'K' implicitly increases co-channel interference as limited numbers of RBs are available for reuse among D2D pairs, during a given session. Therefore to maintain a minimum data rate over cellular and D2D links we have assumed 10 D2D pairs for given scenario. Different path loss models are considered for cellular and D2D users. Each RB is of 180 KHz. Maximum transmit power of a CUE and of DUE is taken as 26 dBm and 13 dBm respectively.

Figure 2 shows the comparative performance of proposed algorithm (OCF-PC) with existing techniques. Optimal Algorithm implies the exhaustive solution for the considered scenario. 'NOCF' refers to non-overlapping coalition formation approach. 'OCF-WPC' refers to overlapping coalition formation game without power control. It is observed that as the number of D2D pairs increases, OCF-PC offers significant improvement in system sum rate as compared to others. Figure.3 shows that as number of cellular user increases the D2D user throughput also increases. This is so because as N increases channels available

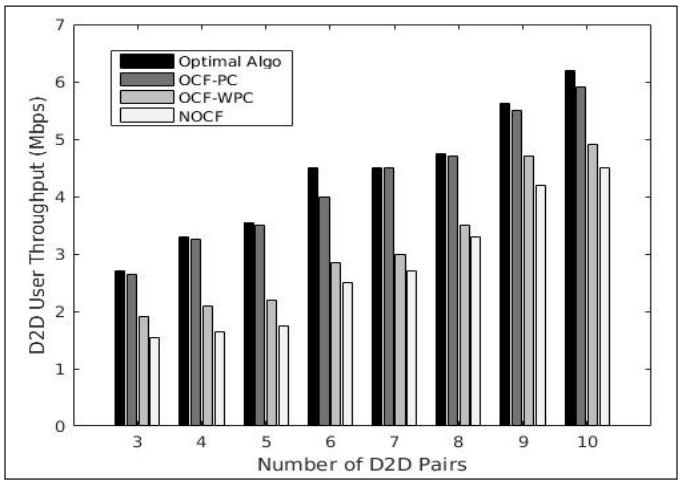


Figure 2. D2D user throughput vs Number of D2D pairs

for sharing also increases as assumed in the system model. This increase is greater in case of OCF-PC because of better interference management and power control as compared to OCF-WPC.

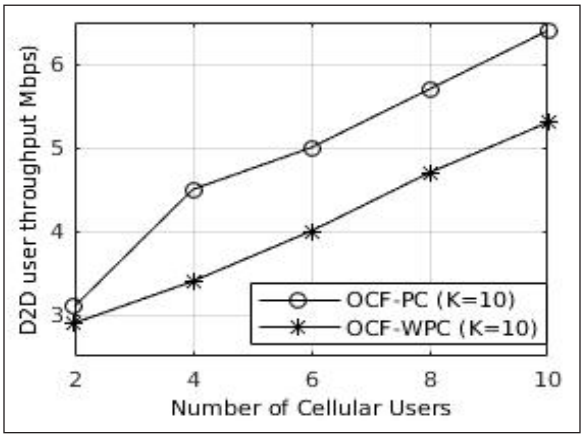


Figure 3. D2D system throughput vs number of cellular users

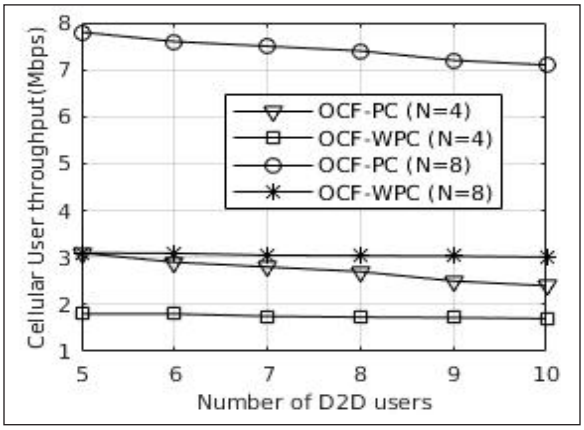


Figure 4. Cellular user throughput vs number of D2D pairs

But increase in number of D2D pairs affects the performance of cellular users because it increases the co-channel interference. The decrease in cellular user throughput can be observed in Figure.4 as 'K' increases. Higher cellular user throughput is achieved in case of OCF-PC for all values of 'N' because of less co-channel interference as compared to OCF-WPC.

Maximum achievable data rate decreases as the distance between source and destination increases. Therefore as the distance between D2D transmitter and receiver increases the D2D user throughput also decreases as shown in Figure.5. The throughput variation is shown for different values of 'N'.

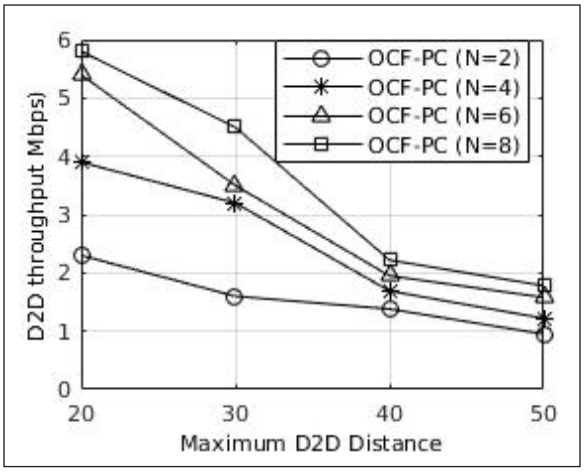


Figure.5. D2D user throughput vs maximum D2D pair distance (in meters)

CONCLUSION

In this work, we have addressed the problem of co-channel interference existing in D2D communication while allow multiple resource allocation for a single D2D pair. Multiple resource allocation and interference based power control scheme is implemented by applying overlapping coalition formation game model and water-filling approach for power distribution over multiple RBs. The proposed scheme provides increased transmission bandwidth and better interference management over each RB, which lowers the impact of co-channel interference. The effectiveness of this scheme is shown in terms of D2D user throughput. D2D user throughput is improved by 46.15% as compared to existing game and matching theoretic techniques. Our future work may consider the effect of co-channel interference in multiple cell scenario while considering the mobility of the D2D users.

REFERENCES

Ciou, S. A., Kao, J. C., Lee, C. Y., & Chen, K. Y. (2015). Multi-sharing resource allocation for device-to-device communication underlaying 5G mobile networks. In *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)* (pp. 1509-1514). Hong Kong, China.

Datsika, E., Antonopoulos, A., Zorba, N., & Erikoukis, C. (2016). Green cooperative device–to–device communication: A social–aware perspective. *IEEE Access, 4*, 3697-3707.

Doppler, K., Rinne, M., Wijting, C., Ribeiro, C. B., & Hugl, K. (2009). Device-to-device communication as an underlay to LTE-advanced networks. *IEEE Communications Magazine*, 47(12), 42-49.

Feng, D., Lu, L., Yuan-Wu, Y., Li, G. Y., Feng, G., & Li, S. (2013). Device-to-device communications underlaying cellular networks. *IEEE Transactions on Communications, 61*(8), 3541-3551.

Fodor, G., Dahlman, E., Mildh, G., Parkvall, S., Reider, N., Miklós, G., & Turányi, Z. (2012). Design aspects of network assisted device-to-device communications. *IEEE Communications Magazine, 50*(3), 170-177.

Janis, P., Chia-Hao, Y. U., Doppler, K., Ribeiro, C., Wijting, C., Klaus, H. U. G. L., & Koivunen, V. (2009). Device-to-device communication underlaying cellular communications systems. *International Journal of Communications, Network and System Sciences*, *2*(03), 169-247.

Jiang, F., Wang, B. C., Sun, C. Y., Liu, Y., & Wang, X. (2016). Resource allocation and dynamic power control for D2D communication underlaying uplink multi-cell networks. *Wireless Networks, 24*(2), 549-563.

Kim, K., Han, Y., & Kim, S. L. (2005). Joint subcarrier and power allocation in uplink OFDMA systems. *IEEE Communications Letters*, *9*(6), 526-528.

Li, Y., Jin, D., Yuan, J., & Han, Z. (2014). Coalitional games for resource allocation in the device-to-device uplink underlaying cellular networks. *IEEE Transactions on Wireless Communications*, *13*(7), 3965-3977.

Li, Y., Wu, T., Hui, P., Jin, D., & Chen, S. (2014). Social-aware D2D communications: Qualitative insights and quantitative analysis. *IEEE Communications Magazine*, *52*(6), 150-158.

Ma, X., Liu, J., & Jiang, H. (2016). Resource allocation for heterogeneous applications with device-to-device communication underlaying cellular networks. *IEEE Journal on Selected Areas in Communications, 34*(1), 15-26.

Rodziewicz, M. (2015). Location-based mode selection and resource allocation in cellular networks with D2D underlay. In *European Wireless 2015; 21th European Wireless Conference; Proceedings of* (pp. 1-6). Budapest, Hungary.

Semiari, O., Saad, W., Valentin, S., Bennis, M., & Poor, H.V. (2015). Context-aware small cell networks: How social metrics improve wireless resource allocation. *IEEE Transactions on Wireless Communications*, *14*(11), 5927-5940.

Sobhi-Givi, S., Khazali, A., Kalbkhani, H., Shayesteh, M. G., & Solouk, V. (2017). Joint mode selection and resource allocation in D2D communication based underlaying cellular networks. *Telecommunication Systems, 67*(1), 47-62

Song, L., Niyato, D., Han, Z., & Hossain, E. (2014). Game-theoretic resource allocation methods for device-to-device communication. *IEEE Wireless Communications, 21*(3), 136-144.

Sun, J., Zhang, T., Liang, X., Zhang, Z., & Chen, Y. (2016). Uplink resource allocation in interference limited area for D2D-based underlaying cellular networks. In *Vehicular Technology Conference (VTC Spring)* (pp. 1-6). Nanjing, China.

Wang, B., Chen, L., Chen, X., Zhang, X., & Yang, D. (2011). Resource allocation optimization for device-to-device communication underlaying cellular networks. *In Vehicular Technology Conference (VTC Spring)* (pp. 1-6). Nanjing, China.

Wang, L., & Tang, H. (2016). Critical Technologies for D2D Communications. In *Device-to-Device Communications in Cellular Networks* (pp. 13-24). Cham: Springer.

Wang, T., Song, L., Saad, W., & Han, Z. (2016). *Overlapping Coalition Formation Games in Wireless Communication Networks*. Cham: Springer.

Wu, X., Tavildar, S., Shakkottai, S., Richardson, T., Li, J., Laroia, R., & Jovicic, A. (2013). FlashLinQ: A synchronous distributed scheduler for peer-to-peer ad hoc networks. *IEEE/ACM Transactions on Networking (TON)*, *21*(4), 1215-1228.

Xia, C., Xu, S., & Kwak, K. S. (2015). Overlapping coalition formation games-based resource allocation for device-to-device communication in LTE-A network. *Transactions on Emerging Telecommunications Technologies, 26*(12), 1295-1305.

Xiao, Y., Chen, K. C., Yuen, C., Han, Z., & DaSilva, L. A. (2015). A Bayesian overlapping coalition formation game for device-to-device spectrum sharing in cellular networks. *IEEE Transactions on Wireless Communications, 14*(7), 4034-4051.

Xu, C., Song, L., Han, Z., Li, D., & Jiao, B. (2012). Resource allocation using a reverse iterative combinatorial auction for device-to-device underlay cellular networks. In *Global Communications Conference (GLOBECOM)* (pp. 4542-4547). Anaheim, CA, USA.

Xu, C., Song, L., & Han, Z. (2014). *Resource management for device-to-device underlay communication* (pp. 1-79). Berlin: Springer.

Xu, S., Xia, C., & Kwak, K. S. (2016). Overlapping coalition formation games based interference coordination for D2D underlaying LTE-A networks. *AEU-International Journal of Electronics and Communications, 70*(2), 204-209.

Zhang, C., Sun, Y., Mo, Y., Zhang, Y., & Bu, S. (2016). Social-aware content downloading for fog radio access networks supported device-to-device communications. In *2016 IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB)* (pp. 1-4). Nanjing, China.

Zhao, J., Liu, Y., Chai, K. K., Chen, Y., & Elkashlan, M. (2017). Many-to-Many Matching With Externalities for Device-to-Device Communications. *IEEE Wireless Communications Letters, 6*(1), 138-141.

Pertanika J. Sci. & Technol. 26 (4): 1881 - 1897 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Experimental Investigations and Exergetic Assessment of 1 kW Solar PV Plant

Ranjana Arora[1]* and Rajesh Arora[2]

[1]*Renewable Energy Department, ASAS, Amity University Haryana, Gurgaon-122413, India*
[2]*Department of Mechanical Engineering, Amity University Haryana, Gurgaon-122413, India*

ABSTRACT

The huge potential of solar energy and increasing demand of energy, made researchers to work on solar photovoltaic systems. In this paper, experimental investigations and exergy based thermodynamic assessment of 1 kW solar photovoltaic (SPV) plant have been carried out. The system is installed at Amity University Gurgaon, India. With the aim to assess the performance/efficiency of the plant, two exergy techniques have been applied based on concepts of thermodynamics and chemical/photonic energy of input solar insolation. The input energy and exergy at different wavelengths ranging from 0.4 μm-0.7 μm have been formulated and illustrated. The electrical and operating parameters of SPV plant includes short-circuit current, open-circuit voltage, temperature of photovoltaic (PV) modules, and fill factor are found, carrying an experiment on a sunny day of 5th October 2017. The variations of electrical exergy input at different fill factors have been computed which signifies its role in characteristic behavior of PV system. The energy/exergy efficiencies are found to be between 7.76% to 9.98% and 9.86% to 11.63% whereas the photonic energy/exergy efficiencies are found to be between 4.85% to 11.24% and 6.08% to 12.89%. It is also found that the temperature of SPV plays a vital role on exergy efficiency and it can be improved with a mechanism which removes the generated heat from the system. With the experimental results, it can be noticed that the exergy loss increases as the temperature of SPV module goes up.

Keywords: Chemical/Photonic energy, energy analysis, exergy analysis, electrical energy, solar insolation rate, solar photovoltaic system, thermal energy

ARTICLE INFO

Article history:
Received: 07 December 2017
Accepted: 11 June 2018
Published: October 2018

E-mail addresses:
ranjana1219@rediffmail.com (Ranjana Arora)
rajesharora1219@rediffmail.com (Rajesh Arora)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Energy demand is increasing day by day and lot of research is going on to explore the sustainable and clean sources of energy. SPV systems are direct energy conversion devices which converts input solar radiations into useful electrical output. The intensity of solar radiations depends on many factors like season, local weather conditions, location of the place and to fully utilize solar energy orientation of the device also plays a very vital role. For Indian perspective, solar energy has got the huge potential as future energy resource, since we have almost 260-300 clear sun days. India (between 6^o to 32^o N latitude) lies in most sunny regions of the world with annual average solar radiations intensity varying between 500 to 800 W/m^2. This is sufficient to fulfil numerous energy demand through solar PV or solar thermal energy conversion routes. In current years, the solar energy technologies are widely acceptable due to steep fall in their investment cost, eco-friendly output and longer life with low maintenance requirements. Solar PV systems are better than the solar thermal as the former needs small installation areas, easy operation/ maintenance tools, variable capital cost of investment and adaptability with other systems like thermoelectric, and solar thermal. Solar PV cells convert input solar radiations (direct/ diffused) into useful electrical output through a semiconductor solid state device. The input solar radiations comprise of direct and diffused components. The diffused radiations are due to scattering, reflection/transmission process of solar radiations in atmosphere by dust, fumes, pollutants or other minute particles. SPV systems work on both direct and diffused radiation components and the maximum output can be obtained when the input rays fall perpendicular to the surface. Therefore, before installation of solar PV power system, side survey needs to be carried out. It gives us the information of input solar intensity, tilt for PV system, wind speed and other meteorological data. A fixed SPV system with proper tilt and appropriate estimation of monthly/yearly average solar radiation intensity can yield maximum output (Duffie & Beckman, 1991). A solar cell is made up of p-n semiconductor as a diode. When solar radiation having photons of high energy falls on solar cell, electrons of n-type combine with holes in p-type. This is due to the recombination process which generates electrical output in the circuit known as electrical exergy. A typical single diode configuration of solar cell is illustrated in Figure 1. It consists of a diode, series/shunt/load resistances namely R_s, R_{sh} and R respectively and a constant current source generating current. The first law efficiency of SPV system is measured as the ratios of power output to the product of solar insolation and the area of the module. The power generated is a product of V_{mp}/I_{mp} (Voltage/Current at peak power points). Energy/Exergy analyses are carried out on the basis of first/second laws of thermodynamics.

Energy analysis cannot give the complete efficiency investigations as it is based only on the quantitative aspects of energy used and the process efficiency. It does not consider the areas where there is reduction in the true energy potentials of the system taking

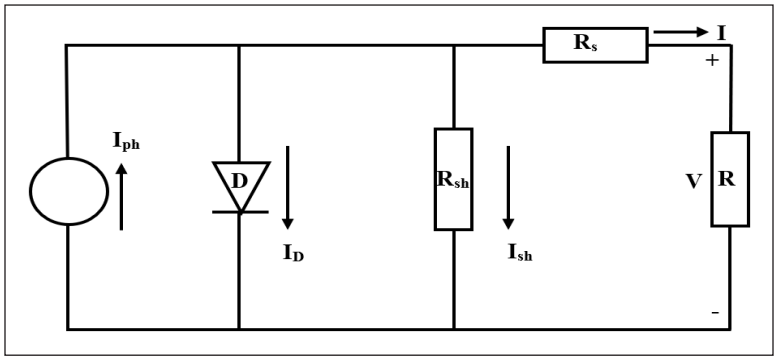


Figure 1. Representation of solar cell in single diode configuration

place. It accounts the amount of available energy to do work. Conversely, the maximum possible useful output attained from energy is known as exergy (Hepabsli, 2008). Exergy analysis presents the true thermodynamic potential and accounts for the usable/unusable (irreversibility) energy fractions. The thermodynamic analysis and economic investigations have been done with the help of exergetic analysis which is proven to be a versatile and prevailing technique (Rosen et al., 2008, 1996, 1988; Rosen, 1999). Likewise, the thermodynamic investigations on SPV systems can be accomplished using exergy analysis as it quantifies the usable and unusable energy known as exergy and irreversibilities respectively. It can also evaluate the efficiency of energy utilization for any conversion process/system and can find the irreversibilities present in the system. With a proper measure, these irreversibilities can be minimized to yield maximum possible output. Therefore, it is gaining wide popularity as it is very helpful in improving the system performance. Several researchers have significantly contributed in the performance characterization of SPV plant. Joshi and Tiwari (2007) enunciated energy/exergy investigations on PV/T parallel plate air collectors under four climatic conditions and formed the range of efficiency variations for the system. Sahin et al. (2007) investigated the thermodynamic assessment of solar modules in context with the exergy analysis for possible system improvement. The comparative analyses of SPV/PV-T systems based on of energy/exergy efficiency have been carried out by Joshi et al. (2009). Akyuz et al. (2012) carried out analytical comparison on simulated and experimental results for a PV system installed in Turkey. Rajoria et al. (2012) investigated the thermal energy/exergy for cascaded SPV array. Afterwards, Saloux et al. (2013) enunciated the electrical/thermal models for the analysis of SPV/PV/T systems to identify the exergy destructions. Sudhakar and Srivastava (2014) performed the energy/exergy investigations for 36 W solar PV modules on a particular day. The experimental investigations on exergetic efficiency of SPV-T water collector system were performed by Yazdanpanahi (2015). Khan and Arsalan (2016) studied concentrated solar power technique and solar photovoltaics to enhance the efficiency/performance of sustainable electrical

output generation technologies. An integrated photovoltaic and unglazed transpired collector has been investigated by Gholampour and Ameri (2016) on the basis of first/second laws of thermodynamics. They emphasized on the significance of energy investigations and electrical/thermal conversion of energy in the whole heat transfer process. Shukla et al. (2016) reviewed PV modules integrated with buildings on the basis of parametric and chemical potential methods. They studied the exergy efficiency for different wavelengths in visible range of spectrum. The comprehensive review on the performance assessment and exergetic analysis of various solar electricity production routes were presented by Bayrak et al. (2017). Afterwards, Ozalp and Bayat (2017) demonstrated various exergy efficiency techniques for thermodynamic investigations of solar PV system. In context with literature review, the major objectives achieved in the present work are as follows:

1. Parametric and photonic energy/exergy of 1 kW installed SPV plant analysed.
2. An accurate model for the prediction of SPV plant performances developed.
3. The thermodynamic considerations for heat transfer study and identification of exergy losses in SPV plant presented.
4. The time domain variations of energy and exergy inputs, efficiencies for the range of wavelengths in the visible region of spectrum investigated.

METHODS

In the present work, the energy/exergy analyses have been carried out for 1kW solar power plant installed at Amity University Gurgaon, as shown in Figure 2. It consists of four 250 W Microsun: MS24250M solar panels connected in electrically parallel configurations. In the given experimental setup, the current from each module is 7.6A and the total current from the system is 30.4A with 44.91V specifications. The system is installed with DC-DC voltage converter and an inverter to suffice the heating and cooling requirement of the building. The required voltage is easily obtained by an installed DC-DC converter and the high current requirement at the application end is obtained by connecting the modules in parallel configurations. The specifications of the 250W module has been mentioned in the Table 1. The longitude and latitude of the place are 76.717^{o} East and 28.15^{o} North. The fluctuations in the ambient temperature are noticed in the span of 4 to 48^{o}C in the whole year at above place. The experimental testing of SPV system has been done and the measurement of the variables viz. I_{sc}, V_{oc}, solar insolation, wind speed and the surrounding temperature etc. has been accomplished between 9 AM to 5 PM with an average gap of 1 hour. Figure 3(a) shows the input solar insolation rate measuring device, called pyranometer (Eppley Laboratory Model, PSP36851F3) whereas the lux meter, as shown in Figure 3(b), has been utilized to measure the ambient temperature and moreover, the solar insolation

has also been measured/verified with it due to its easy portability and robustness. Figures 3(c) shows the anemometer that is used for the measurement of wind speed. The input parameters of the chosen module have been mentioned in the Table 2.

Table 1
Input parameters

Input parameter (s)	Value
Nominal operating cell temperature (NOCT)	47^{o}C
Stefan's Boltzman Constant	5.67×10^{-8} $Wm^{-2}K^{-1}$
Emissivity of the panel	0.89
Sun Temperature	5780 K

Figure 2. Experimental test of the thermoelectric module

Table 2
Specifications of 250 W Photovoltaic module

Model	Micro Sun: MS24250M
Maximum Power	250 W
Open Circuit Voltage	44.91 V
Short Circuit Current	7.6 A
V_{mp}	37.030 V
I_{mp}	6.910 V
No. of cells	72
Dimensions	1666 mm × 978 mm × 42 mm
Weight	18.65 kg
Fill Factor	0.75

(a) (b) (c)

Figure 3. (a) Pyranometer; (b) Lux meter; and (c) Anemometer

Thermodynamic Assessment

Energy Analysis. According to the law of conservation of energy, the equations on energy balance for a system can be expressed as

$$\sum_{in} en_{in} m_{in} - \sum_{out} en_{out} m_{out} + \sum_{k} Q_k - W = 0 \quad (1)$$

Where en_{in}/en_{out} represents the inlet and outlet energies, m_{in}/m_{out} are the mass flow rates through inlet/outlet, Q_k is heat flow rate whereas W is the work done across the system.

Further, according to law of conservation of mass,

$$m_{in} = m_{out} \quad (2)$$

The steady state assumption for the system gives $m_{in}=m_{out}=0$. Therefore, Eqn (1) can be rewritten as

$$\sum_{k} Q_k - W = 0 \quad (3)$$

The conversion of input solar insolation into useful electrical output in SPV system can be calculated by the current generated in the system given as

$$I = I_1 - I_0 \times \exp^{q\times(V-IR_s)/A\times K\times T} \quad (4)$$

The output electrical power can be represented as

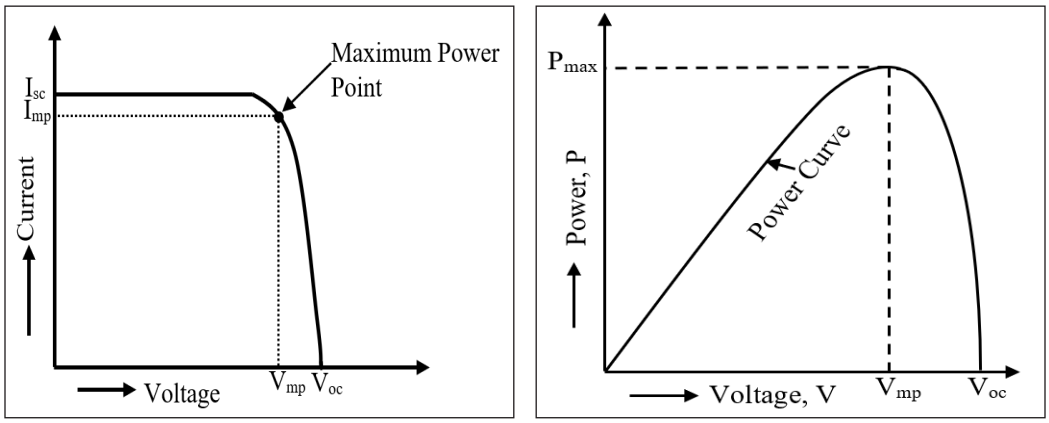
$$P_{el} = I \times V \quad (5)$$

The maximum magnitude of output electrical power for the system is represented as

$$P_{\max} = V_{oc} \times I_{sc} \times FF = V_{mp} \times I_{mp} \quad (6)$$

The efficiency of solar cell is a measurement of its capability to translate input solar radiations into useful electrical output where the output electrical power comprises of the

current/voltage from the system. This efficiency of conversion varies with the intensity of solar radiations. Besides, at the maximum power point, one can obtain the maximum conversion efficiency from the system with the help of maximum voltage V_{mp} and current I_{mp} as shown in Figure 4.



(a) (b)

Figure 4 . (a) I-V Characteristics and (b) P-V characteristics of solar photovoltaic system

The output criterion followed in the system is $V_{mp}<V_{oc}$ and I_{mp} is also smaller than the magnitude of short circuit current I_{sc}. The fill factor restricts the maximum power obtained from the system and represents it by maximum power point.

$$FF = \frac{V_{mp} \times I_{mp}}{V_{oc} \times I_{sc}} \tag{7}$$

The conversion energy efficiency of SPV plant is computed with the help of the following formula

$$\eta_{en} = \frac{V_{mp} \times I_{mp}}{A \times E} \tag{8}$$

Here, A is the area of the module in m^2 and E is solar insolation in Wm^{-2}.

The input solar radiations being absorbed the solar cell get converted into useful electrical output and thermal energy, which further get dissipated by the phenomenon of conduction, convection and radiation. However, the design and geometry of SPV module effects the heat transfer process rate. In order to obtain the SPV system efficiency, T_c, ahe oerating temperature, should be computed and assumed to be homogenous on the exposed surface. This temperature varies with the surrounding temperature. As the surface

temperature increases, it can cause significant dip in the system efficiency and some cooling phenomenon should be employed at the back side to improve the efficiency. The cooling can be air or water depending upon the application.

Conventional Method. The conversion of input solar radiations into useful electrical output by solar PV plant is known as electrical exergy which can be entirely exploited for the suitable purpose. On the other hand, the thermal output cannot contribute towards the productive yield and appears as high heat loss in the system.

Under steady state assumption of an open system, the equations on exergy balance is given as

$$\sum_{in} ex_{in} m_{in} - \sum_{out} ex_{out} m_{out} + \sum_{k} Ex^{Q} - Ex^{W} - I^{'} = 0 \tag{9}$$

Where, ex_{in}/ex_{out} denotes the input and output exergies for the system, m_{in}/m_{out} are the mass flow rates through the system at inlet/outlet, E_x^Q is exergetic heat flow rate whereas E_x^W is the net exergetic work done across the system. The term I', denotes the electrical and thermal exergy losses in the system.

Further, the steady state assumption for the system gives $m_{in}=m_{out}=0$. Therefore, Eqn (9) can be rewritten as

$$\sum_{k} Ex^{Q} - Ex^{W} - I^{'} = 0 \tag{10}$$

Where,

$$I^{'} = T_{amb} S_{gen} \tag{11}$$

The input exergy for SPV system including the effects of solar input radiations intensity, area of the panel and ambient temperature can be computed as

$$Ex_{in} = A \times E \left\{ 1 - \frac{4}{3}\left(\frac{T_{amb}}{T_{sun}} \right) + \frac{1}{3}\left(\frac{T_{amb}}{T_{sun}} \right)^{4} \right\} \tag{12}$$

The electrical and thermal exergy output for the system can be formulated as

$$Ex_{el} = V_{oc} \times I_{sc} \times FF \tag{13}$$

$$Ex_{th} = Q \left\{ 1 - \left(\frac{T_{amb}}{T_{\mathrm{m}l}} \right) \right\} \tag{14}$$

For the net exergy output calculations, the module temperature can be evaluated with respect to normal operating cell temperature (NOCT) as

$$T_{ml} = T_{amb} + (NOCT - 20) \times (E/800) \tag{15}$$

The system experiences the heat loss which includes the convective and the radiative factors and can be represented as

$$Q = A \times (h_c + h_r) \times (T_{ml} - T_{amb}) \quad (16)$$

Where, h_c and h_r are convective and radiative heat loss coefficients respectively and are represented as

$$h_c = 2.8 + 3v \quad (17)$$

$$h_r = \varepsilon\sigma(T_{sky} + T_{ml})(T_{sky}^2 + T_{ml}^2) \quad (18)$$

For further calculations, the effective sky temperature is

$$T_{sky} = T_{amb} - 6 \quad (19)$$

Based on the above formulation, the net exergetic efficiency for the system is

$$\eta_{ex} = \frac{Ex_{out}}{Ex_{in}} \quad (20)$$

Photonic Method. The input solar radiations consist of tiny energy packets called photons. On the basis of their energy level, the energy of solar radiations can be calculated, known as energy of photons. This can be computed as

$$E_{ph}(\lambda) = \frac{hc}{\lambda} \quad (21)$$

It can be calculated by the estimation of rate of number of photons falling on a particular surface. According to Markvart (2003), there are $4.4X10^{17}$ number of photons which are falling every second on $1cm^2$ area on the earth surface on a hazy day. With the assumption of the magnitude of solar constant as 1367 W/m^2, this figure can further be formulated with the multiplication of solar radiation intensity with the factor $4.4X10^{21}/1367$. As a result, the photonic energy for SPV system is the multiplication of figure of photons which falls every second on a $1m^2$ area, photonic energy and the area absorbing solar energy. Henceforth, this is the estimation of available chemical potential or photonic energy of any photovoltaic system and can be formulated by the given formulae (Marti & Luque, 2004):

$$E_{cp} = E_{ph}(\lambda)\left\{1 - \frac{T_{ml}}{T_{sun}}\right\} \quad (22)$$

$$= AN_{ph}\left(\frac{hc}{\lambda}\right)\left\{1 - \frac{T_{ml}}{T_{sun}}\right\} \quad (23)$$

This can further lead to the photonic exergy of the system if multiplied with the energy conversion efficiency of solar cell and is computed as

$$Ex_{cp} = \eta_{en} \times A \times N_{ph} \left(\frac{hc}{\lambda} \right) \left\{ 1 - \frac{T_{ml}}{T_{sun}} \right\} \tag{24}$$

RESULTS AND DISCUSSION

The real-time analysis has been carried out for a typical day of October 2017 at Amity University Gurgaon, Haryana, India and the data is utilized to evaluate the influence of environmental conditions on the throughput of 1 kW SPV system. On the basis of the thermodynamic second laws and the available energy of solar radiation photon, the energy/exergy inputs have been evaluated and presented. The analytical comparisons among photonic energy/exergy efficiencies have been computed and presented with respect to various input parameters. It is noticed that exergetic efficiency represents better perspective of the system performance as it takes into the account all losses and irreversibilities in the system. Therefore, it is highly effectual and powerful tool for evaluation and investigations of systems in real world. Additionally, the root means square percent deviation' factor, p and correlation linear coefficients, q can be formulated with the help of given expressions as

$$p = \sqrt{\frac{\sum (p_i)^2}{n}} \tag{25}$$

$$\text{Where, } p_i = \left(\frac{Y_{1(i)} - Y_{2(i)}}{Y_{1(i)}} \right) \times 100 \tag{26}$$

$$q = \frac{N\left(\sum Y_2 Y_1\right) - \left(\sum Y_2\right)\left(\sum Y_1\right)}{\sqrt{N(\sum Y_2^2) - (\sum Y_2)^2} \times \sqrt{N(\sum Y_1^2) - (\sum Y_1)^2}} \tag{27}$$

This needs to be done for the comparison of various computed efficiencies with the predicted ones. The magnitude of q varies between -1 and 1 and is a measure of direction/intensity of linear relation amongst two chosen variables. If the value of q approaches 1, then the two variables are assumed to have strong positive relationship on linear curve and *vice-versa*.

Figure 5 illustrates the variations of input solar insolation and ambient temperature with time between 9AM and 5PM. The variation in wind speed is observed between 0.98-1.2 m/s which influences the convective coefficient of heat transfer between surface of SPV array and the surroundings. The ambient temperature is observed to be varying between 307 K to 310.2 K on the chosen day. The input solar insolation intensity observes the variation between the range 710 W/m^2 and 1150 W/m^2.

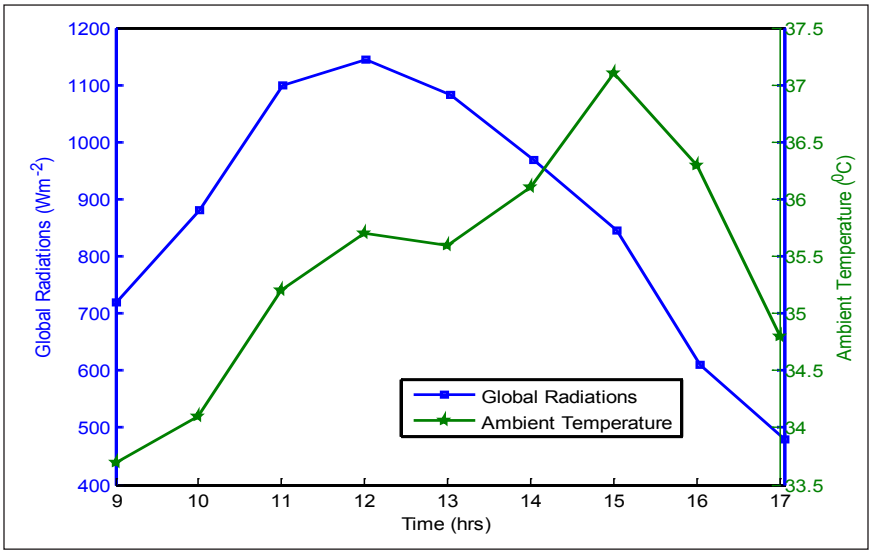


Figure 5. Variation of input solar insolation and surrounding temperature with time

The energy inputs at different values of wavelengths in visible region (0.4 µm-0.7 µm) are plotted with respect to time in Figure 6. Both solar radiations and energy inputs follow the same pattern with time. It is observed that input energy and photonic energy for the PV plate possess a synchronism with respect to root mean square percentage deviation (p) with 13.76% and correlation linear coefficient (q) as 0.978 for a typical chosen wavelength of 0.55 µm. However, p/q approaches their lowest magnitudes (p=5.36% and q=0.988) at a wavelength of 0.60 µm.

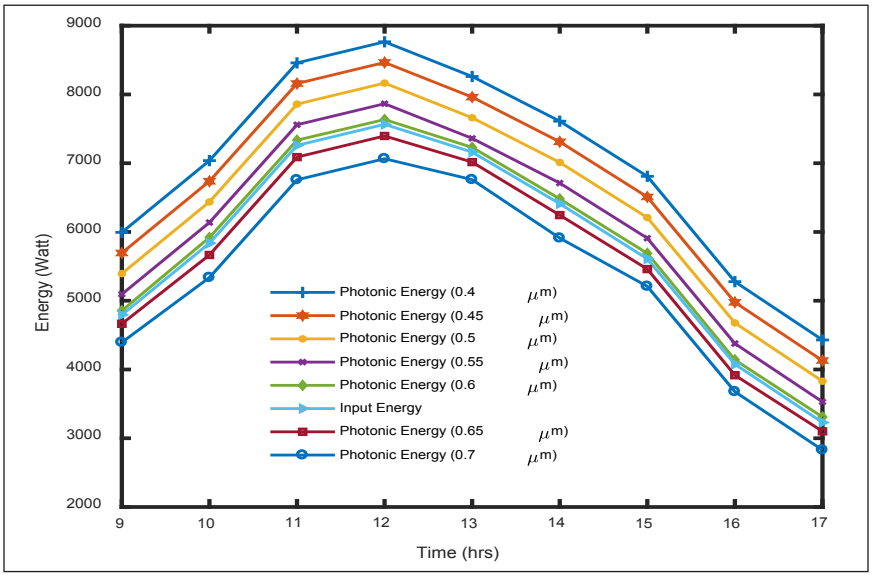


Figure 6. Variations of exergy input and exergy efficiency with time

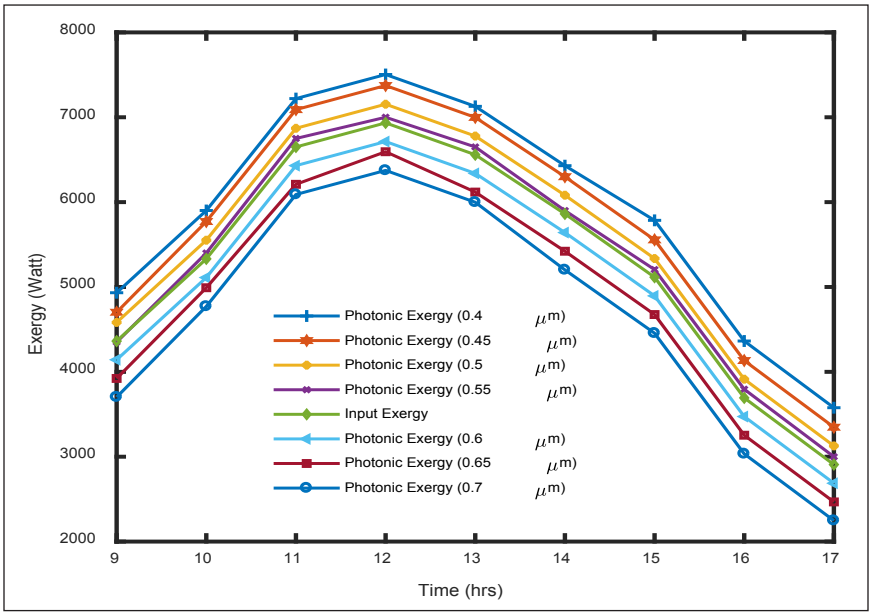


Figure 7. Exergy Inputs at different wavelengths in visible region on 5th October, 2017

On the other hand, the exergy inputs at different values of wavelengths in visible region (0.4μm-0.7μm) are plotted with respect to time in Figure 7. It is noticed that input exergy and photonic exergy (0.55 μm wavelength) are in agreement with p=4.98% and q=0.994. However, at maximum solar intensity, the exergy inputs approach their peak values due to larger value of sun temperature.

The variations in electrical exergy for different fill factors with respect to time has been illustrated in Figure 8. It signifies that the output electrical exergy can be considerably enhanced by improving the fill factor in SPV system. It can be achieved by minimizing the optical and thermal losses in the system and thus improving the throughput across it.

On the basis of experimental observations, it has been found that the loss of exergy in the whole day is exceeding more than 85% and has been shown in Figure 9. Large amount of useful exergy is lost due to inefficiency of silicon modules to convert input solar radiations input useful output. The complete analysis yields an average exergy/energy efficiency of 10.25% and 8.4% respectively. The system possesses low exergetic efficiency due to inferior quality of energy output and suffers huge losses in useful exergy during the conversion process. The magnitude of exergy destruction factor is more than 88% in the system and needs lot more of improvement in material aspect.

Figure 10 illustrates the variations in the magnitude of energy/exergy efficiencies along with the photonic/chemical potential on a typical day of October 2017. It clearly shows that the two efficiencies changes with the temperature and to attain their maximum value,

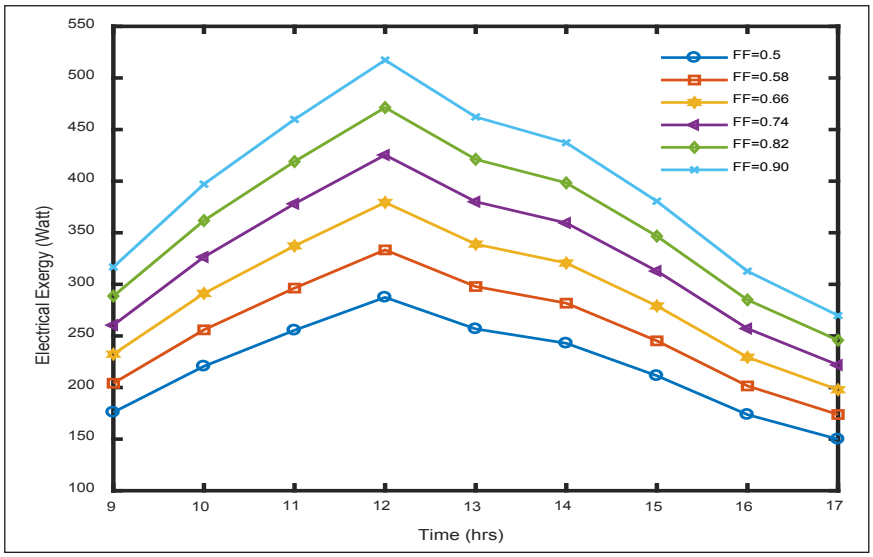


Figure 8. Variations of electrical exergy and input solar exergy with time

the temperature of SPV module should work closer to the operating temperature of the cell. The system observes low energy/exergy efficiency for the whole day with reference to ideal 100% reversible phenomenon. This is because of irreversible nature of solar PV energy conversion process. Only a certain part of input solar radiations gets converted into useful electrical output. Rest get wasted into heat and other optical losses. The conventional available silicon cell modules suffer huge exergy loss due to low conversion efficiency. However, it increases with the rise in input solar insolation.

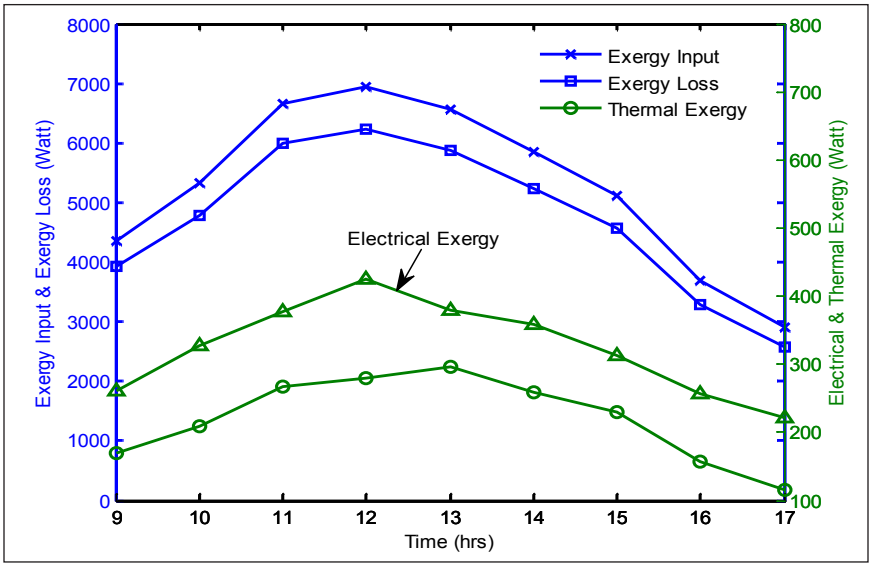


Figure 9. Variations of thermal, electrical and exergy loss with time

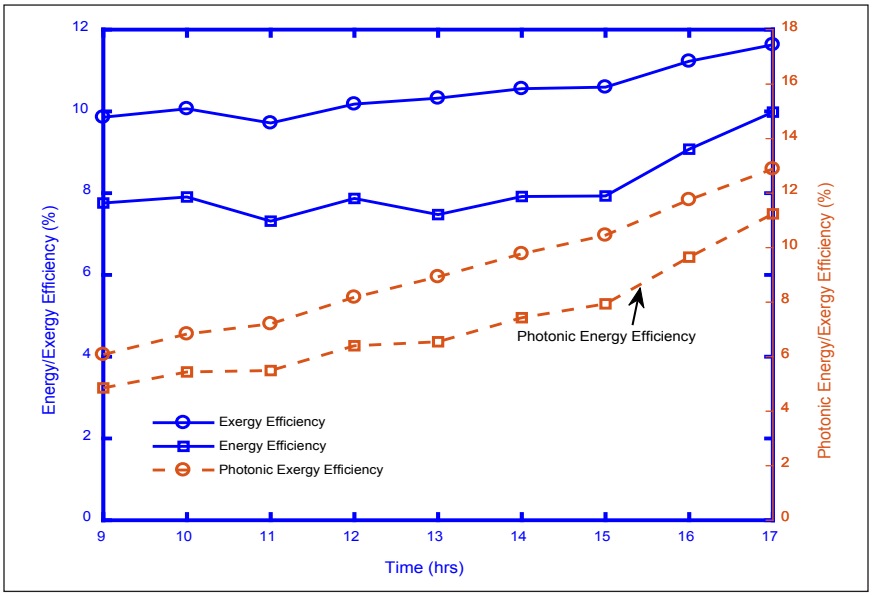


Figure 10. Energy/exergy efficiency versus time

CONCLUSION

The parametric and photonic energy/exergy investigations have been carried out for performance estimation of 1 kW solar photovoltaic system installed at Amity University, Gurgaon, India. Two different methods for the assessment of energy/exergy analysis are demonstrated for real time monitoring of the SPV system on a typical day of October 2017. The obtained experimental figures are utilized to investigate an optimal temperature in order to get maximum exergy efficiency. Various exergetic losses are also quantified for energy conversion process. The major outcomes of the research work are summarized as below:

1. With the study, it is found that exergetic analysis based on both thermodynamic principles and photonic/chemical potential yield realistic and time bound results. Thus, either of the method can be employed for the performance assessment of thermodynamic modeling, installation and planning of SPV plants.
2. It is also observed that fill factor is the key feature for the estimation of exergetic efficiency of SPV plant and projects a clearer perspective for possible improvement in the system. High value of fill factor leads to better exergetic efficiency.
3. The exergetic efficiency of the system rises with input solar radiation and attains its maximum value. Thereafter, it shows a slight dip and get stabilized. The system possesses maximum energy/exergy efficiencies of 11.24% and 12.89% respectively. However, the low exergetic efficiency signifies that the modules are unable to process high energy content of available solar radiations to useful electrical output.

4. The exergetic efficiency possesses an inverse relationship with the ambient temperatures. The reason is the rise in module temperature and various irreversibilities occurring in the system, whereas useful electrical output increases with increase in intensity of input solar radiations.
5. The visible spectrum wavelength is a key factor in estimating photonic/chemical potential energy/exergy. It has been observed that at lower wavelengths, the energy/exergy possess higher values.
6. The current study suggests the development of better and low-cost materials for solar PV energy conversion process. This can significantly improve system efficiency.

It is observed that SPV systems are direct energy conversion devices and can yield huge amount of power at minimum operating/maintenance cost. Moreover, these eco-friendly and environmentally sustainable sources of energy can solve all the power problems in coming era. Additionally, new optimization technique for exploring design and performance parameters of solar panels can lead to better results for the system. Moreover, focused investigations on modelling and designing are required in context with better efficiency and output performance of the SPV modules. The work can be extended to get the optimal efficiency by suitable formulations and computations of the system parameters. The experiment can further be conducted for more number of days in a year and research findings can be precisely verified.

ACKNOWLEDGEMENT

The authors would like to thank Amity University Haryana-122413, India for providing support in carrying out this valuable research work.

REFERENCES

Akyuz, E., Coskun, C., Oktay, Z., & Dinçer, I. (2012). A novel approach for estimating of photovoltaic exergy efficiency. *Energy, 44*(1), 1059–1066.

Bayrak, F., Abu-Hamdeh, N., Alnefaie, K. A., & Oztop, H. F. (2017). A review on exergy analysis of solar electricity production. *Renewable and Sustainable Energy Reviews, 74*, 755-770.

Duffie, J. A., & Beckman, W. A. (1991). *Solar Engineering Thermal Processes* (2nd ed.). New York: Wiley.

Gholampour, M., & Ameri, M. (2016). Energy and exergy analyses of photovoltaic/thermal flat transpired collectors: experimental and theoretical study. *Applied Energy, 164*, 837–856.

Hepabsli, A. (2008). A key review on exergetic analysis and assessment of renewable energy resources for a sustainable future. *Renewable and Sustainable Energy Reviews, 12*(3), 593-661.

Joshi, A. S., & Tiwari, A. (2007). Energy and exergy efficiencies of a hybrid photovoltaic thermal (PV/T) air collector. *Renewable Energy, 32*(13), 2223-2241.

Joshi, A. S., Dincer, I., & Reddy, B. V. (2009). Thermodynamic assessment of photovoltaic systems. *Solar Energy, 83*(8), 1139-1149.

Khan, J., & Arsalan, M. H. (2016). Solar power technologies for sustainable electricity generator–a review. *Renewable and Sustainable Energy Reviews, 55*, 414–425.

Markvart, T. (2003). *Solar Electricity* (2nd ed.). Chichester, U.K: John Wiley & Sons Ltd. Marti, A., & Luque, A. (2004). *Next generation photovoltaics high efficiency through full spectrum utilization*. Bristol, U.K: Institute of Physics Publishing Ltd.

Ozalp, M., & Bayat, M. A. (2017). Comparative thermodynamic analysis on different exergetic efficiency methods for a solar photovoltaic module. *International Journal of Exergy, 24*(2-4), 325-343.

Rajoria, C. S., Agrawal, S., & Tiwari, G. N. (2012). Overall thermal energy and exergy analysis of hybrid photovoltaic thermal array. *Solar Energy, 86*(5), 1531-1538.

Rosen, M. A. (1999). Second law analysis approaches and implications. *International Journal of Energy Research, 23*(5), 415-429.

Rosen, M. A., & Hooper, F. C. (1996). Second law analysis of thermal energy storage system. In *Proceedings of First Trabzon International Energy and Environment Symposium* (pp. 361-371). Trabzon, Turkey.

Rosen, M. A., Dincer, I., & Kanoglu, M. (2008). Role of exergy in increasing efficiency and sustainability and reducing environmental impact. *Energy Policy, 36*(1), 128–137.

Rosen, M. A., Hooper, F. C., & Barbaris, L. N. (1988). Exergy Analysis for the Evaluation of the Performance of Closed Thermal Energy Storage Systems. *Transactions on ASME: Journal of Solar Energy Engineering, 110*(4), 255-261.

Sahin, A. D., Dincer, I., & Rosen, M. A. (2007). Thermodynamic analysis of solar photovoltaic cell systems. *Solar Energy Materials & Solar Cells, 91*(2-3), 153-159.

Saloux, E., Teyssedou, A., & Sorin, M. (2013). Analysis of photovoltaic (PV) and photovoltaic /thermal (PV/T) systems using the exergy method. *Energy Building, 67*, 275-285.

Shukla, A. K., Sudhakar, K., & Baredar, P. (2016). Exergetic assessment of BIPV module using parametric and photonic energy methods: A review. *Energy and Buildings, 119*, 62-73.

Sudhakar, K., & Srivastava, T. (2014). Energy and exergy analysis of 36 W solar photovoltaic module. *International Journal of Ambient Energy, 35*(1), 51-57.

Xu, C., Wang, Z., Li, X., & Sun, F. (2011). Energy and exergy analysis of solar power tower plants. *Applied Thermal Engineering, 31*(17-18), 3904–3913.Yazdanpanahi, J., Sarhaddi, F., & Adeli, M. M. (2015). Experimental investigation of exergy efficiency of a solar photovoltaic thermal (PVT) water collector based on exergy losses. *Solar Energy, 118*, 197-208.

APPENDIX

Nomenclature

A	Area (m^2)
c	Speed of light ($3X10^8$ ms^{-1})
E	Solar insolation (Wm^{-2})
Ex	Exergy rate (W)
En	Energy rate (W)
F	Fill Factor
h	Plank Constant ($6.625X10^{-34}$Js)
h_c	Convective heat transfer coefficient ($Wm^{-2}K^{-1}$)
h_r	radiative heat transfer coefficient ($Wm^{-2}K^{-1}$)
I	Current (A)
I'	Exergy consumption (W)
N_{ph}	Number of photons/$s.cm^{-2}$
NOCT	Nominal operating cell temperature (^oC)
P	Power (W)
R_s	Series resistance
R_{sh}	Shunt resistance
T	Temperature (K)
U	Overall heat transfer coefficient ($Wm^{-2}K^{-1}$)
v	Wind velocity (ms^{-1})

Greek symbols

η	Efficiency
σ	Stefan's Boltzman Constant ($Wm^{-2}K^{-4}$)
ε	Emissivity
λ	Wavelength (m)

Subscripts

amb	ambient
el	electrical
th	thermal
in	inlet
ex	exergy
en	energy
out	outlet
ph	photonic
max	maximum
ml	module
oc	Open circuit
mp	Maximum power
sc	Short circuit
cp	Chemical potential

Superscripts

Q	Heat
W	Work

Abbreviations

PV	Photovoltaic
SPV	Solar Photovoltaic
NOCE	Normal operating cell temperature

Pertanika J. Sci. & Technol. 26 (4): 1899 - 1911 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Longitudinal Changes in Suspended Sediment Loading and Sediment Budget of Merbok River Catchment, Kedah, Malaysia

Ismail, W. R.[1]*, Ibrahim, M. N.[1] and Najib, S. A. M.[1,2]

[1]Section of Geography, School of Humanities, Universiti Sains Malaysia (USM), 11800 Minden, Penang, Malaysia

[2]Department of Geography & Environment, Faculty of Human Science, Universiti Pendidikan Sultan Idris (UPSI), 35900 Tanjong Malim, Perak, Malaysia

ABSTRACT

Merbok river catchment situated in the Kedah State receives its input from Bongkok River and Puntar River flowing down and joining Lalang River to flow down to the Merbok Estuary. The Merbok catchment (440 km^2) is experiencing several degrees of complex land uses activities that poses some impact on the suspended sediment production of the Merbok river. A study was conducted to investigate the suspended sediment loading of rivers draining the Merbok catchment from January to December 2013. Suspended sediment budget of the Merbok catchment were estimated. The river suspended sediment concentrations (SSC) and suspended sediment (SS) load increased during wet season compared to dry season. The SS loads increases from upper catchment to river mouth. The sediment loadings were divided into three segments- the upstream, middle segment and lower segment. The SS loads increased from 10 t yr^{-1} in the upper part of Bongkok river to 3336 t yr^{-1} in upper segment. The sediment loading then increase to 4299 t yr^{-1} in the middle segment of the catchment (at Bongkok 4), and then exiting the Merbok Estuary, as the lower segment, with a total amount of sediment output estimated at 7156 t yr^{-1}. From this total sediment output, most of the sediment source came from the tributaries; the Bongkok River at B3 (3337 t yr^{-1}), Puntar River (2924 t yr^{-1}) and Lalang River (1370 t yr^{-1}), which were much higher than its proportion in terms of its length and drainage area. As a conclusion, the inconsistence in SSC in the river were influenced by the various anthropogenic activities (especially agriculture and urbanization activities) in

ARTICLE INFO

Article history:
Received: 16 January 2018
Accepted: 28 May 2018
Published: 24 October 2018

E-mail addresses:
wruslan@usm.my (Ismail, W.R.)
nazrulibrahim@yahoo.com (Ibrahim, M.N.)
sumayyah@fsk.upsi.edu.my (Najib, S.A.M.)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

the catchment area which necessitate future land use and sediment control to avoid sediment and possible nutrient loading into the estuary.

Keywords: Kedah, land use changes, Merbok River catchment, sediment load, sediment budget

INTRODUCTION

Sediment eroded from slope land, bare land and agricultural lands, all ended up in the rivers draining the whole of the world's landscape. Sediments play an important role in the river ecosystem where they are responsible for transporting many materials and contaminants to the river (Ismail, 1996). Sediment ended up in water bodies could also brought damages to water quality when eroded soil enters surface waters. The characterization of suspended sediment transport in rivers is difficult due to the rapid of human exploitation and unpredictable natural hydrologic events. Excessive soil loss is linked to lack of the sustainable use of natural resources, land exploitation and degradation and poor land management.

Concerns regarding the effects of human activities on degradation in river systems, erosion and sedimentation have frequently been raised. High sediment loads can, in particular, result in major problems for human such as affecting flood risk and boat traffic (Sheffield et al., 1995); decrease of water quality and water supply, increasing the cost of treating water and other related purposes (Walling, 2009). High sediment suspended in the river can kill or irritate fish gills and suffocates organisms if significant concentrations and durations occur (Bash et al., 2001).

A sediment budget is an application of the continuity equation and it can be regarded as a form of geomorphic accountancy (Slaymaker, 2003). It is also considered as a quantitative inventory of all the sediment inputs, outputs and storage within a defined system. An understanding of the sources of sediment delivered to, stored within, and exported from, an estuary is important for a number of environmental issues including maintenance of navigational channels, light availability for primary productivity, reduction of dissolved oxygen concentrations and the transport and accumulation of particle-bound nutrients and contaminants and their eventual transport to the continental shelf (Balls, 1992; Cloern, 1987; Eyre & McConchie, 1993).

The sediment production is expected to increase in the near future due to rapid economic activities which could lead to the widespread changes in erosion rates, sediment flux and sediment loads transported by the world's rivers shown the high-frequency of human activities and climate change. The amount of the sediment load transported by rivers in the catchment area has a very important implication for biology (Wood & Armitage, 1997), geomorphology (Brandt, 2000) and hydrological system (Syvitski, 2003). Additional, the sediment input from several catchment is important in the management of rivers and water resources, particularly in the context of integrated water resources for domestic

use, ecotourism, and biodiversity of rivers, hydrology and hydraulics (O'Keeffe, 2009). The sediment loads may influence on crop productivity and food security, changes in nutrient cycling and flux of many key elements and nutrients, water quality, river channel morphology, delta development and also the aquatic ecosystems and numerous of habitats supported by the river (Walling, 2009). The sediment loads can be estimated by developing sediment budget. The budget is suitable for identifying and accounting of the sediment movement, into and out of a site on the catchment. Developing sediment budget is an attempt to identify the sources, sinks and pathways of eroded material content within catchment area (Slaymaker, 2003). For many years, sediment budget has been developed at scales ranging from small catchment to a large catchment (Owens et al., 1997; Walling et al., 2002).

This paper aims to estimate a simple sediment budget in the Merbok catchment area in Kedah, and to assess the various pathways and sediment sources from various land uses and tributaries in the catchment. This study also investigates the role of seasonal effect and the impact of human activities on variations in the concentration of SSC and the amount of input/output of suspended-sediment (SS) load.

MATERIALS AND METHODS

Study Area

The investigation focused on the Merbok River Catchment area from Gurun area down to the Merbok river at Kampong Sg. Lalang, Kedah within the latitude within latitudes and 5° 39' N to 5° 41' N and longitudes 100° 20' E to 100° 24' E, respectively (Figure 1). The catchment area of Merbok river is 440 km^2. Figure 1 shows all 15 sampling sites in the Merbok river catchment chosen in this research based on relation between the main river and its tributaries. The river is classified into the upper, middle and lower segment longitudinally according to the dominant gradient of the river (Rosgen 1994), where the uppermost segment called segment A+ should have a slope of more than 10%; followed by A between 4-10%, B between 2-4%; C having less than 2% and so on. Table 1 shows the river classification for Bongkok and Merbok river.

Land Uses

Land use in the Merbok river catchment (440 km^2) is very complex. The catchment area is experiencing several degrees of complex land uses activities ranging from some major activities for example oil palm (24%) and rubber (9.7%) in the upper and middle part of the catchment. Some part of the middle and lower part of the catchment is urbanized with urban and residential land uses (22%), and the lower part of the catchment is dominated by swamp and mangrove (9.4%). Other types of land uses in the Merbok catchment area

includes, forest (6.7%) which is diminishing; and some industries (5.5%) in areas like Gurun, Bakar Arang and Sungai Petani, and some paddy (5.9%) and mixed horticulture (3%) in the lower part of the catchment area (Figure 2) (Malaysia, 2010).

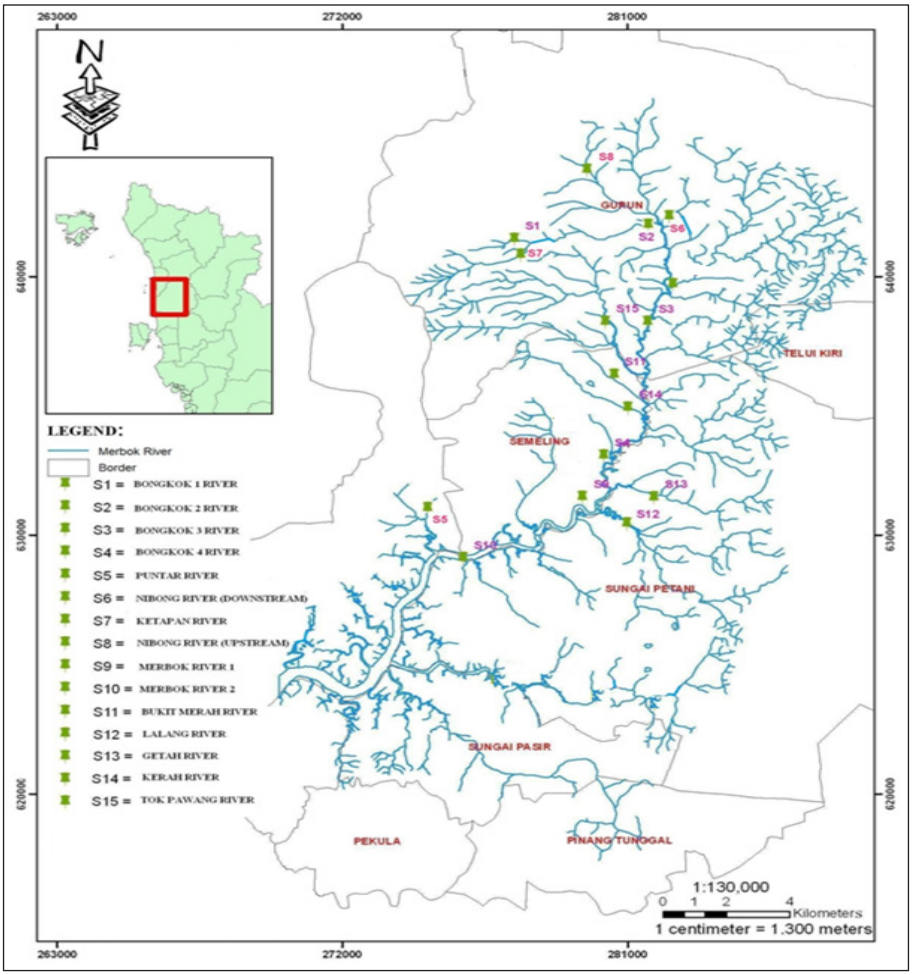


Figure 1. Location of all the sampling stations from the upper Merbok to the estuary

Table 1
The distance, gradient and classification of the river reach of the main river Bongkok and Merbok

Segment	River Reach	Distance (km)	Gradient (%)	Classification
Upper	Bongkok 1	2.05	20.46	A+
	Bongkok 2	4.42	18.56	A+
	Bongkok 3	7.27	11.28	A+
Middle	Bongkok 4	14.15	5.54	A
Lower	Merbok 1	18.9	3.08	B
	Merbok 2	26.65	1.16	C

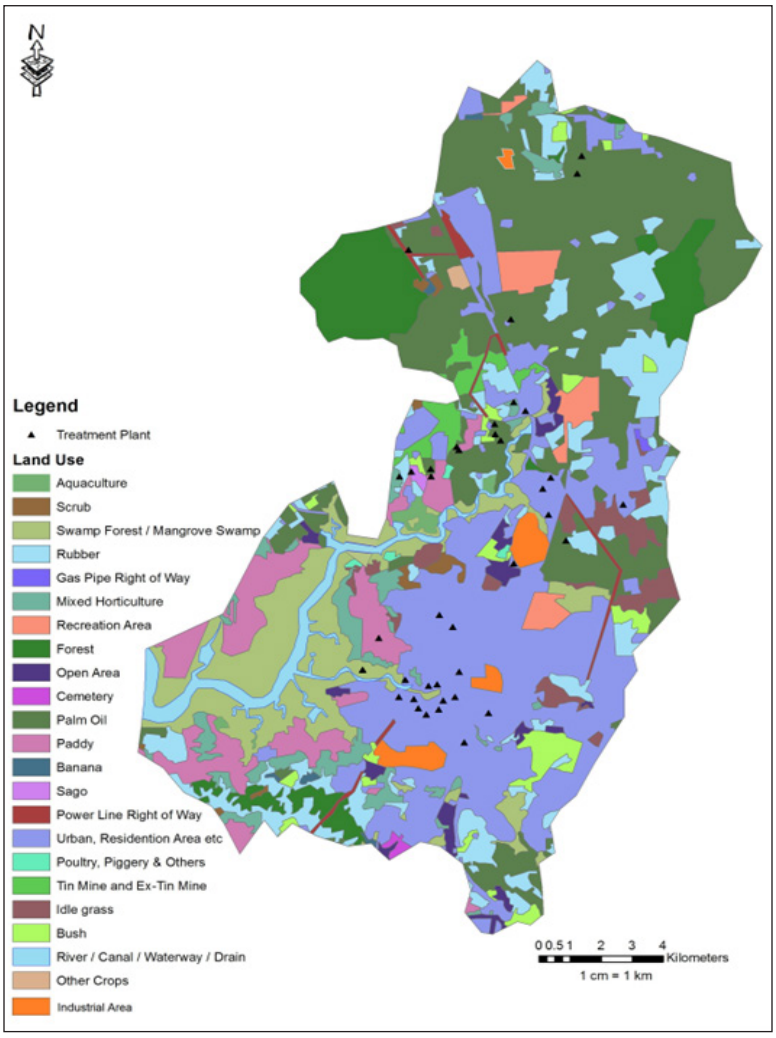


Figure 2. The land uses in the catchment area of Merbok (Malaysia, 2010)

Climate

Rainfall data were recorded at two rainfall stations in the catchment area namely the Jerai Peak and Jambatan Syed Omar. Jerai Peak recorded range of rainfall from 65 mm - 536 mm in five months period from September 2011 to February 2012 (Malaysia, 2012) shown in Figure 3. Jerai Peak or Gunung Jerai is a major water source flowing into the Merbok River. Jambatan Syed Omar Station recorded a little lower rainfall range within the same period (61 mm- 435 mm). Figure 3 shows that the rainfall patterns in the catchment area are divided into dry season in the months of September, December and January, and wet months occur in October, November and February.

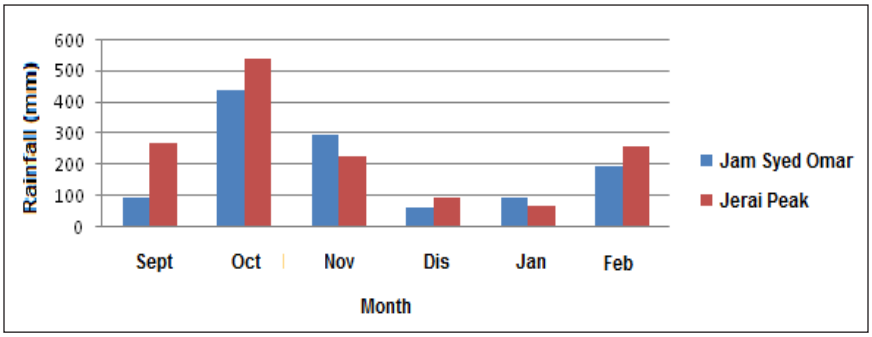


Figure 3. Rainfall in the Merbok catchment recorded at two stations; Jerai Peak and Jam Syed Omar in between September 2011 until February 2012

The Sampling Protocol

The water quality and physical parameters sampling was carried out every fortnight at selected river cross sections (15 sites) throughout the catchment area from January to December 2013. Figure 4 shows five main river cross sections for the main tributary the Bongkok River. The river width ranging from 3 m at B1; 10-15 m at B2 and B3; 40 m at B4 and nearly 60 m at M1.

Water samples were collected in three replicates representing a river cross-section segments and preserved in a polyethylene bottle. The water sampling procedure was done together with an *in situ* water quality monitoring scheme using YSI 556 multi-parameter. The water samples were collected at depths of about 0.5 m from water surface, and at the center of each segment, after dividing the river width into 3 equal segments. Water samples were transported in an ice-cooled container to the Hydrology laboratory, School of Humanities, Universiti Sains Malaysia for further analysis.

River discharge which is the volume of water that flows past a certain point in a stream over a specific period of time, was calculated based on the velocity area method (Shaw et al., 2010) and the calculation of SS loading was based on Littlewood (1992).

Determination of Suspended Sediment Concentration (SSC)

The determination SSC was carried out using filtration methods using membrane Glass Fiber Filter ADVANTEC 47mm. After filtration, the membrane filter was placed in a petri dish for drying in an oven at 105ºC and was left overnight. The membrane filter was cooled at room temperature before being weighed for it final weight. SSC in the sample was calculated by using an equation following APHA (1992). To obtain an estimate of SSC, the difference between weight of filter + dried residue and weight of filter was calculated (APHA, 1992).

Sediment Loading Calculation

The sediment loading is based on the calculating the product of suspended sediment concentration (SSC) with river flow (Q) in a given period. The total amount of suspended load (in tonnes/year) was then calculated by multiplying the weighed sediment concentrations (SSC) in mg/L with the stream discharge (m^3s^{-1}) with an interval of 14 days.

To calculate the sediment loading, the average sample load estimation has been used. The load of sediments transported through a river cross-section during a time interval using calculation described by Littlewood (1992);

$$L = K\left(1/n\sum_{i=l}^{n} Q\,(i).C(\mathrm{i})\right)$$

where:

L = load

K = the period of load estimation

C = sample concentration

Q = discharge

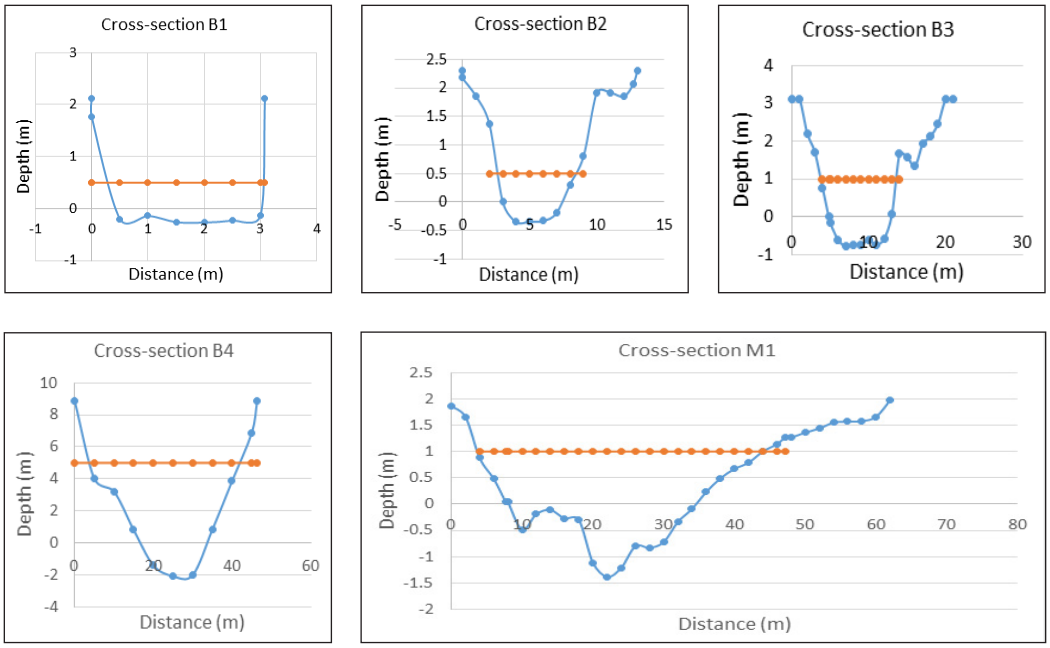


Figure 4. Five main river cross sections from the upper segment B1 to the lower segment at M1. The river cross section is shown by the blue lines, while red lines are the water depth at the time of measuring the cross sections.

The accuracy of estimating the sediment load depended upon the availability of the river flow and concentration data at a sufficiently high frequency (relative to the variation in flow and concentration during the period of estimation) (Littlewood, 1992).

RESULTS

The Suspended Sediment Load

The result in Table 2 shows the sediment loading of each station and were placed into three segments of the catchment i.e the upstream segment, middle segment and the lower segment towards the river estuary (Table 2). The main river of the Merbok catchment originated as Bongkok river as the upper, middle and part of the third segment, then it became the Merbok river. So the longitudinal flow was from Bongkok 1 (B1) to Bongkok 4 (B4) and Merbok 1 (M1) (Table 2).

The total SS loads increased from 10 t yr^{-1} in the upper part of B1 to 261 t at B2. This is about 25 fold increases in the SS loads. Then the SS loads further increases downstream to 3337 t yr^{-1} at Bongkok 3 (B3). This represents an increase of only 12 fold between B2 and B3. The sediment loading then increase to 4299 t yr^{-1}at Bongkok 4 (B4). The increase in SS load was only about 0.3 fold but later it increases to 0.66 fold at Merbok 1, where the sediment loading was estimated at 7156 t yr^{-1}. Based on the total SS load, Merbok River catchment was experiencing a lot of sediment removal. Most of the sediment source came from the tributaries like the Puntar River and Lalang River, which was much higher than its proportion in terms of its length and drainage area.

In wet season, the SS loads increased, with a much higher increases were noted in the upper segment (34 and 26 fold, respectively in Table 2). However, the wet period loading decreased slightly (0.25 fold) between B2 and B3, but later increased 0.8 fold between B4 and M1. In the dry season however the SS loading was much lower due to the lower rainfall activity that drive the hydrological processes in the catchment. The SS loading was very little ranging from 12 fold in the upper segment, 9 and 0.13 fold between B2 and B3, and 1.6 fold in the lower segment (Table 2).

The overall SS loading increased, from the upper to lowest segment, by almost 700 fold for total SS load ranging from 400 fold in the dry period to about 1250 fold increases in the wet period. The SS load was higher at M1 because it received all the sediment from the upper tributaries.

The overall loading between wet and dry period was not much different which was only about 26 tonnes (3.5%). However, there was very significant difference in SS loading between the wet and dry period observed at B3 (123%) and B4 (48%). These two stations were located in the middle part of the catchments which is highly active in the human activities and land use changes. With a characteristics of great seasons variations and precipitation in an alternation in wet period and dry period, the annual sediment load from

Table 2
Summary of suspended sediment (SS) loading of Merbok River catchment at the main channel of Bongkok and Merbok River in Wet and Dry Season

Catchment Segment	Sites	Total Loads (t yr^{-1})	Wet Season (t)	Dry Season (t)	Increase in total SS load (fold)	Increase SS load in wet season (fold)	Increase in SS load in dry season (fold)
	Bongkok 1	10	0.6	1.8			
	Bongkok 2	261	20.8	23.4			
(B2-B1)/B1					25.1	33.7	12.0
	Puntar	2924	469.7	206.5			
Upper	Ketapan	242	34.8	19.7			
Segment	Nibong 1	384	40.2	30.7			
(S1)	Nibong 2	483	61.8	43.9			
	Bongkok 3	3337	557.0	248.7			
(B3-B2)/B2					11.8	25.8	9.63
	Bukit Merah	208	23.6	14.2			
Middle	Getah	310	68.6	24.3			
Segment	Kerah	241	33.8	13.0			
(S2)	Tok Pawang	284	45.5	25.4			
	Bongkok 4	4299	415.6	280.7			
(B4-B3)/B3					0.29	-0.25	0.13
	Lalang	1370	1442.7	115.1			
Lower Segment	Merbok 1	7156	748.5	722.6			
(M1-B4)/B4					0.66	0.80	1.57
(M1-B1)/B1					714.6	1246.5	400.4

the Merbok River catchment discharged to sea showed a very high SS load (ca. 7000 t/year) exported to the river mouth. However, this amount is small compared to Sg Kurau (51, 270 t/year) feeding into Bukit Merah reservoir (Ismail and Najib, 2011); but comparable to Sg. Jarum (5,400 t/year) and Pelarit River (11,400 t/year) reported by Rahaman and Ismail (2010) reaching into Timah Tasoh lake which is also flowing through agricultural land use activities.

DISCUSSION

Naturally, the sediment is transported from land surfaces and slopes to river channel during rainy season. The sediment sources may originate from sources such as landslides in cleared areas (Collison and Anderson, 1996); slope collapses along rivers (Madej, 2001); raindrop impact and subsequent surface wash of exposed soils on roads, agriculture area and open

area with poorly compacted soil (Ziegler & Giambelluca, 1997). The understanding of the pattern of sediment delivery and its pathways related to hydrological processes is needed in these investigations of sediment budget. It is very difficult to develop prudent long-term management plans for limitation and reduce the sediment budget in large catchments (Walling, 1983).

Human disturbance, such as agriculture and urbanization activity, makes sediment available for transport during rainfall events (Ismail and Hashim, 2014). The SS load trend showed an increase during wet season but decline during dry season. This is due to the increase sediment mobility during rainy season when the surface runoff increases, transporting more sediment (Ismail, 1996; Sarria-Villa et al., 2016). The annual SS load also showed an increasing trend from the upper segment to lower segment. However, the rate of increases in sediment loading slightly reduced between the middle to lower segment by about 29%, but toward the lower end of the segment of the catchment, the percentage of sediment load increased by 66%, suggesting a higher sediment loading happened in the lower part of the catchment where a lot of urbanization was taking place. The increase in human activities for example a lot of housing and industrial development takes place in the middle and lower part of the catchment.

Changes in the sediment load of a river can give rise to numerous problems. Erosion and sediment transport processes are sensitive due to their close links to climate, wide range of human activities, land cover, land use and the hydrology of a river catchment (Walling, 2009). This reflects a combination of factors. Human activities that affect the river ecosystem by sedimentation and higher suspended solids from forest cutting and land-clearance, the expansion of agriculture, land use practices, urbanization and infrastructural development and dam construction (Walling & Fang, 2005). These impacts have been especially severe in river ecosystems. Ziegler et al. (2000) found that surface erosion rates from agricultural roads, the main source of sediment to the river was more than eight times higher than from footpaths and agricultural lands.

In future, the variability in sediment loading necessitates the need for long-term monitoring of sediment yield. River management is imperative to promote harmonious relationship between humans and nature in reducing sediment input to the river systems. Sediment budgeting approach facilitate in identifying the sediment sources urgently needed attention. The variations of discharge flow and runoff due to seasonal factors shows difference in SS load related to sediment concentration during the wet season period compared to that during dry season period. Other than that, SSC content in the river can also be affected by the influence of anthropogenic activities in the catchment area.

CONCLUSION

This study shows that the sediment load increases from the upper segment of river catchment (B1) which has very little disturbances in terms of land use activities as compared to many land use activities in the catchment, between B3, B4 and M1. The downstream changes in SS loads and increasing loading suggest that sediment is being added to the main rivers from various tributaries in the catchment. The total suspended sediment loading was around 7000 t yr^{-1} and was comparable to other Malaysian river fluxes associated with agricultural activities. The various land uses in the catchment areas contributed to the varying SS loads as shown by the results. The loading in wet period was almost 50% and more than the dry period loads especially in the middle segment of the catchment due to varying land uses in this region suggesting the effect of rainfall on the eroding the land surfaces with varying land cover. The inconsistencies in SSC in the river were influenced by the various anthropogenic activities (especially agriculture and urbanization activities) in the catchment area which necessitate future land use and sediment control to avoid sediment and possible nutrient loading into the estuary.

ACKNOWLEDGEMENTS

The authors would like to thank Mohamad Khushairy Zahir and Mohamad Adam Omar for assisntant during the field sampling. We acknowledged Universiti Sains Malaysia for RU Grant no. 1001/PHUMANITI/816229.

REFERENCES

APHA. (1992). *Standard Methods for the Examination of Water and Wastewater* (18th ed.). Washington, DC: American Public Health Association.

Balls, P. W. (1992). Nutrient behaviour in two contrasting Scottish estuaries, the Forth and Tay. *Oceanologica Acta, 15*(3), 261–277.

Bash, J., Berman, C. H., & Bolton, S. (2001). *Effects of Turbidity and Suspended Solids on Salmonids*. Seattle, United States: University of Washington Water Center.

Brandt, S. A. (2000). Classification of geomorphological effects downstream of dams. *Catena, 40*(4), 375–401.

Cloern, J. E. (1987). Turbidity as a control on phytoplankton biomass and productivity in estuaries. *Continental Shelf Research, 7*(11-12), 1367–1381.

Collison, A. J. C., & Anderson, M. G. (1996). Using a combined slope hydrology/stability model to identify suitable conditions for landslide prevention by vegetation in the humid tropics. *Earth Surface Processes and Landforms, 21*(8), 737-747.

Eyre, B. D., & McConchie, D. (1993). The implications of sedimentological studies for environmental pollution assessment and management: examples from fluvial systems in north Queensland and Western Australia. *Sedimentary Geology, 85*(1-4), 235–252.

Ismail, W. R. (1996). The role of tropical storms in the catchment sediment removal. *Journal of Bioscience, 7*(2), 153-168.

Ismail, W. R., & Hashim, M. (2014). Changing trends in rainfall and sediment fluxes of Kinta River, Perak, Malaysia. In *Proceedings of a symposium on Sediment Dynamics from the Summit to the Sea* (pp. 340-345). New Orleans, Louisiana, USA.

Ismail, W. R., & Najib, S. A. M. (2011). Sediment and nutrient Balance of Bukit Merah Reservoir, Perak (Malaysia). *Lakes and Reservoirs: Research and Management, 16*(3), 179-184.

Littlewood, I. G. (1992). *Estimating Contaminant Loads in Rivers: A Review*. Institute of Hydrology Report No.17, Institute of Hydrology, Wallingford, UK.

Madej, M. A. (2001). Erosion and sediment delivery following removal of forest roads. *Earth Surface Processes and Landforms, 26*(2), 175-190.

Malaysia (2010). Department of Agriculture Malaysia, 2010. Putrajaya.

Malaysia (2012). Drainage and Irrigation Department, 2012. Kedah.

Owens, P. N., Walling, D. E., He, Q., Shanahan, J. O., & Foster, I. D. (1997). The use of caesium-137 measurements to establish a sediment budget for the Start catchment, Devon, UK. *Hydrological Sciences Journal, 42*(3), 405-423.

O'Keeffe, J. (2009). Sustaining river ecosystems: Balancing use and protection. *Progress in Physical Geography, 33*(3), 339–357.

Rahaman, Z. A., & Ismail, W. R. (2010). Spatial and temporal variation of sedimentation rate in the Timah Tasoh water reservoir, Perlis, Malaysia. *Annals of Warsaw University of Life Sciences- SGGW, Land reclamation, 42*(1), 127-138.

Rosgen, D. L. (1994). A classification of natural rivers. *Catena, 22*(3), 169-199.

Sarria-Villa, R., Ocampo-Duque, W., Paez, M., & Schuhmacher, M. (2016). Presence of PAHs in water and sediments of the Colombian Cauca River during heavy rain episodes, and implications for risk assessment. *Science of the Total Environment, 540*(1), 455-465.

Shaw, E. M., Beven, K. J., Chappell, N. A., & Lamb, R. (2010). *Hydrology in Practice*. Abingdon: Spon Press, Taylor & Francis.

Sheffield, A. T., Healy, T. R., & McGlone, M. S. (1995). Infilling rates of a steepland catchment estuary, Whangamata, New Zealand. *Journal of Coastal Research, 11*(4), 1294-1308.

Slaymaker, O. (2003). The sediment budget as conceptual framework and management tool. In *The Interactions between Sediments and Water* (pp. 71-82). Dordrecht: Springer.

Syvitski, J. P. M. (2003). Supply and flux of sediment along hydrological pathways: research for the 21st century. *Global and Planetary Change, 39*(1-2), 1-11.

Vorosmarty, C. J., Meybeck, M., Fekete, B., Sharma, K., Green, P., & Syvitski, J. P. M. (2003). Anthropogenic sediment retention: major global impact from registered river impoundments. *Global and Planetary Change, 39*(1-2), 169-190.

Walling, D. E. (1983). The sediment delivery problem. *Journal of Hydrology, 65*(1-3), 209–237.

Walling, D. E. (2009). *The impact of global change on erosion and sediment transport by rivers: Current progress and future challenges*. Paris: UNESCO.

Walling, D. E., & Fang, D. (2003). Recent trends in the suspended sediment loads of the world's rivers. *Global and Planetary Change, 39*(1), 111-126.

Walling, D. E., Russell, M. A., Hodgkinson, R. A., & Zhang, Y. (2002). Establishing sediment budgets for two small lowland agricultural catchments in the UK. *Catena, 47*(4), 323-353.

Wood, P. J., & Armitage, P. D. (1997). Biological effects of fine sediment in the lotic environment. *Environmental Management, 21*(2), 203-217.

Ziegler, A. D., & Giambelluca, T. W. (1997) Importance of rural roads as source areas for runoff in mountainous areas of northern Thailand. *Journal of Hydrology, 196*(1-4), 204-229.

Ziegler, A. D., Sutherland, R. A., & Giambelluca, T. W. (2000). Runoff generation and sediment production on unpaved roads, footpaths and agricultural land surfaces in northern Thailand. *Earth Surface Processes and Landforms, 25*(5), 519-534.

Pertanika J. Sci. & Technol. 26 (4): 1913 - 1925 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

WBMFC: Efficient and Secure Storage of Genomic Data

Sanjeev Kumar*, Suneeta Agarwal and Ranvijay

Department of Computer Science and Engineering, MNNIT Allahabad, India

ABSTRACT

With the development of next-generation sequencing technology, a massive amount of genomic data are being generated day by day. To efficiently handle these data for storage, processing and transmission, some specialized genomic data compression techniques are need of today. In the near future, personalized genomics may come into existence where doctors may give the treatment on the basis of patient genome. It creates a huge challenge to securely store and transmit the genomic data over the cloud servers or remote servers. This problem can be solved by applying a combination of encryption and compression techniques. Most of the state of the art algorithms for secure and efficient storage of genomic data adopt the policy of encryption after compression. The computational costs of these algorithms are very high, so there is a need to develop a unified encryption-compression algorithm (encryption during compression) to provide the confidentiality/secrecy also to genomic data. In this paper an approach applying encryption during compression is proposed to efficiently and securely store the genomic data in fasta/multi-fasta file format. Here MWBTC (Modified Word Based Tag Code) and Delta Encoding are used for compression and AES-256 is used for encryption. Experiments show that the proposed algorithm (WBMFC) outperforms the state of the art algorithms in terms of processing time and compression ratio both.

Keywords: Compression, decompression, encryption, fasta, multi-fasta

INTRODUCTION

Genome data is a collection of genetic information (in the form of DNA sequence) of a living organism. Size of such data is very large. For example, one human genome contains 3.2 billion DNA base pairs, which takes 3 GB in memory (Danek, & Deorowicz, 2018). With the advancement of sequencing machine technology huge

ARTICLE INFO

Article history:
Received: 16 January 2018
Accepted: 28 May 2018
Published: 24 October 2018

E-mail addresses:
skgcp86@gmail.com (Sanjeev kumar)
suneeta@mnnit.ac.in (Suneeta Agarwal)
ranvijay@mnnit.ac.in (Ranvijay)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

volumes of genomic data are being deposited in public repositories (NCBI, DDBJ, EBI) and cloud servers for the purpose of research, forensic and diagnosis (Hosseini et al., 2016). The exponential growth of these data creates a severe challenge for secure storage, fast processing, and transmission. Compression with encryption is a crucial tool to address these challenges (Wiese et. al., 2018). It reduces storage space and processing cost along with security and also speeds up data transmission. To store and transmit genomic data efficiently and securely, a unified compression-encryption algorithm is the need of today.

Most of the public repositories store the genomic data in fasta/multi-fasta file format. Fasta is a text based file format to represent a genomic sequence (Danek & Deorowicz, 2018). Fasta file always starts with ">" followed by the header which contains some descriptions about the genome and sequencing machine used, next is the sequence data which contains the characters A, C, G, T, and N. A Multi-fasta file contains several fasta files. Formats of Fasta and Multi-fasta files are shown in Figure 1 and Figure 2 respectively (Sardaraz et al., 2014).

Genomic data has some special characteristics such as large no of repeats (tandem & palindrome) and less no of nucleotides (A/ C/ G/ T) (Dix et al., 2006). To compress the genomic data, general purpose compression algorithms (Gzip, Bzip2, LZMA) do not take

```
>HUMAN_NM_018998_01
TGAGTAGGGCTGGCAGAGCTGGNNNNNNGGCCTCATGGCTGTGTAGTAGC
AGGCCCCCGCCCCGCGACCTGGCCAGGCGATCACTACAGCCGCCCCTGCC
GAACAGNNNNNNNNCCCCATGAGGACCCAGAACCCTCAATGGAGAAGAG
TCAGGATTTGCTGTGCTGCCAGAGTGAACTGGCCTGGTAATTACCCTGCA
GCCTTTCTGGAACAGNNNNNNNNGGTGCTCACAGCCCCAGAGACACCACT
GAGGTAGGAAGCTGCCCTGGAGTGATGTCCTTGGGGCATTGGACAGGGAC
CCTCACCGTAGCCCTCCCTGCAG
```

Figure 1. Fasta file

```
>HUMAN_NM_018998_1
TGAGTAGGGCTGGCAGAGNNNNCTGGGGCCTCATGGCTGTGTAGTAGCAG
GCCCCCGCCCCGCGACCTGGCCAGGCGATCACTACAGCCGCCCCTGCCGA
>Mouse_NM_013908_2
CCCCATGAGGACCCAGAACCCTCAATGGAGAAGAGTCAGGATTTGCTGTG
CTGCCAGAGTGAACTGGCCTNNNNNGGTAATTACCCTGCAGCCTTTCTGG
>HUMAN_NM_018998_3
GTGAGCACGGGCGGCGGGTTGACCCNNNNNTGCCCCCGCCCCACGCCGAC
AGCCTGTCCAGCCCCGGCCTCCCCACAG
>Mouse_NM_013908_4
GTAAGTGTGGGCATTGGGTTGGGCTACCTGTCCCATTGTGCCCTGCCAGCA
GTCTGCCCAGCTGTGGCCTTCCCCCCAG
>HUMAN_NM_018998_5
GGTGCTCACAGCCCCAGAGACACCACTGAGGTAGGAAGCTGCCCTGGAGT
GATGTCCTTGGGGCATTGGACAGGGACCCTCACCGTAGCCCTCCCTGCAG
```

Figure 2. Multi-fasta file

into account the characteristics of these data, while domain specific algorithms (BIND, Deliminate, MFCompress, SeqCompress and Cryfa) utilize the characteristics of these data (palindrome repeats, tandem repeats and less no of nucleotides) (Kumar et al., 2015). For security if required these algorithms use some external tools for encryption. Most of the existing DPCA adopt two level techniques to compress and encrypt the genomic data, i.e., encryption after compression (Figure 4). These techniques take large amount of time to compress & encrypt the genomic data. Thus to reduce the processing time there is a need to develop a unified encryption-compression technique (encryption during compression) (Figure 5).

There is a number of state of the art algorithms to compress the fasta/multi-fasta files. BIND (Sardaraz et al., 2014) uses two binary streams for compression. In the first stream, both A & T are assigned bit 0, and C & G are assigned bit1. In the second stream, both C and T are assigned bit 0, and A & G are assigned bit1. These two streams are independently compressed with LZMA general purpose compression algorithm. The average compression ratio of BIND is 4.3.

DELIMINATE (Mohammed et al., 2012) uses delta encoding for two most frequent characters, remaining characters are encoded with 0 and 1. Average compression ratio is 4.65. Compression time and decompression time are same as of BIND.

MFCompress (Pinho & Pratas, 2013) uses finite text models for encoding the fasta and multi-fasta files. It is based on a probabilistic model that determines probability distribution by calculating the probability of next nucleotide in the genome sequence based on k-previous nucleotides (order-k context). MFCompress uses single finite context model for encoding of header data and order-k context model for encoding sequence data. Compression ratio is same as of Deliminate, but compression and decompression speed are less in comparison to Deliminate.

SeqCompress (Sardaraz et al., 2014) uses arithmetic coding and statistical model. The statistical model is based on the frequency of fragments in the input sequence, decides whether to use fragment-based compression or binary compression. It is a two-pass algorithm. An average compression ratio of SeqCompress is 4.92 but compression time and decompression time are higher than that of BIND and Deliminate.

Cryfa (Pratas et al., 2017) first uses three-bit packing technique to reduce the size of genomic data, and thereafter encryption is applied to this packed data. The compression ratio of Cryfa is same as of DELIMINATE. It also provides encryption.

Existing state of the arts techniques for compression of fasta/multi-fasta files are either dictionary based or statistical-based. Dictionary-based techniques work in two phases: first, a dictionary is created, and then a substitution based method is used to encode the sequence. Such technique requires the large size of dictionary during decompression, it creates the problem of storage and transmission (Darok et. al., 2017). Statistical models

used probabilistic method to predict the next character from the past occurrences of the characters. Such techniques require huge memory during compression and decompression. For security existing tools use two-level techniques, i.e., encryption after compression. The computational cost of such a technique is very high (Jahaan et al, 2017). So there is need to develop a unified technique (encryption during compression) for compressing the fasta/ multi-fasta files. In this paper, a new approach to compress & encrypt the fasta/multi-fasta file is presented. The proposed approach uses unified encryption approach (encryption during compression) to encrypt and compress the genomic data. Proposed approach first divides the fasta/multi-fasta file into two streams: header stream and sequence stream. These two streams are compressed individually with the appropriate compression algorithms. Delta difference encoding is used to compress the header stream while MWBTC (Modified Word Based Tag Code) (Gupta & Agarwal, 2008) is used to compress the sequence stream. Encryption is applied only on the dictionary created by MWBTC at the time of encoding instead of applying it on the whole genome data. Since the size of a dictionary in MWBTC is very small, therefore whole genomic data is encrypted very fast.

METHODS

Proposed method first separates the genomic data (fasta / multifasta) in two streams: sequence stream (W1) and header stream (W2) (files). Stream W2 contains header parts of the input genome sequence while stream W1 contains remaining part of the input genome sequence. If input file is a Multi–fasta file, then along with header data lengths of the corresponding sequences are also stored in W2. Files W2 and W1 are compressed with delta difference encoding and MWBTC respectively as shown in Figure 3. Encryption is applied on the dictionary created by MWBTC to secure the whole genome data.

Details of compression and encryption method used are as follows:

Header Stream Compression

To compress the header stream Delta difference encoding method is used. Finally, it is archived by 7ZIP general purpose compression as shown in Figure 3.

Sequence Stream Compression

To compress the sequence stream all occurrences of non ACGT character N (if present) are first extracted. All positions of N are recorded in a separate file W3. Generally in genomic sequence, N's are present in clusters. Positions of first occurrence in each cluster along with the size of that cluster are stored in W3 as shown in step 1.1 of Example 1. Thereafter remaining characters A, C, G & T of the sequence stream W1 are encoded with the modified word based tag code (MWBTC). MWBTC reads the file W1 and segments it into words of size 4, maximum 256 words are possible. Thereafter frequencies of all words of

W1 are calculated and stored in column "count" of table1. A new table "tempvocab" is formed with three columns: index, word, and count. Index column is for indices of each word. Word column contains all the possible words (A, C, G, and T) of size four stored in lexicographical order. Count columns contain corresponding frequencies of the words obtained from Table 1. A new Table 3 is formed by sorting full rows of Table 2 with respect to the contents of the column count. First column (index) of Table 3 is renamed as shuffled index here. A new column "index" is also added for indices (0-255) as the first column of Table 3.

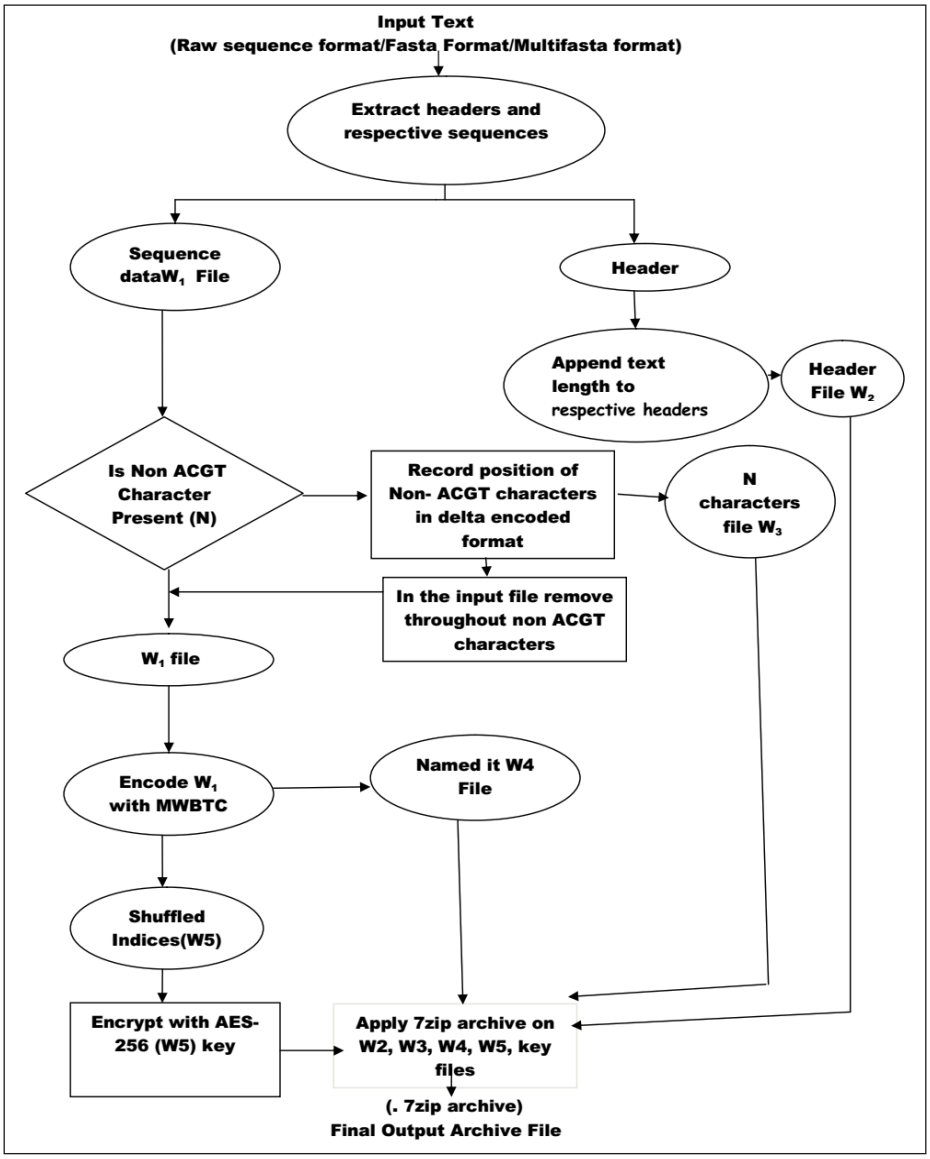


Figure 3. Flow diagram of the proposed approach

Codes are assigned to each word of Table 3 using MEncode method (Algorithm1) and stored them in a new column (fifth column) codes. MWBTC allocate less number of bits for frequents words and more bits for lesser frequent words. Contents of file W1 are encoded using the generated codes in Table 3 and named it as W4 file. The complete encoding process of MWBTC is explained in step 1.2 of Example 1.

Example 1: Let the given genome sequence for secure and efficient storage be:
T = A A A A C G G A A T T N N N N N N G A T A A A A A N N
N N C G A C A A T T G A T A A A A A A A A A A C A C
The size of T is 54 characters.
Steps of compression and decompression by proposed approach are as follows:

Step 1. Compression Process
\\ **Input:** Genome Sequence T
\\ **Output:** Output.7zip - Compressed genome sequence along with secret key and a dictionary.

Step 1.1 (Extraction of non ACGT character N): Input sequence "T" has two clusters of N's at positions 13-18 & 27-30. All N's (if present) are extracted from input genome sequence T. They are stored as 13, 6, 27 and 4 in a new file W3. Here numbers at odd positions represent starting positions of clusters and numbers at even positions represent the size of the respective clusters.

Step 1.2 (Encoding): After extraction of non ACGT character N from input sequence T, resulting input sequence T'=A A A A C C G G A A T T G A T A A A A A C G A C A A T T G A T A A A A A A A A A A C A C
The size of T' is now 44 characters. The content of "vocab" table is presented in Table 1**.** The compressed form of sequence T'' (W4 file):
01110110000101111010000101010010

Contents of the second column of Table 3 are stored in a new file W5. W5 alone is sufficient to decode the above compressed sequence T''. To enhance the security of whole genome sequence it is proposed to encrypt the contents of W5 only using AES-256 (Mahajan & Sachdeva, 2013).

Table 1
Contents of table "vocab"

Index	Word	Count
0	AAAA	4
1	CCGG	1
2	AATT	2
3	GATA	2
4	CGAC	1
5	ACAC	1

Table 2
Contents of table "tempvocab"

Index	Word	Count
0	AAAA	4
1	AATT	2
2	ACAC	1
3	CCGG	1
4	CGAC	1
5	GATA	2

Table 3
Contents of table "finalvocab"

Index	Shuffle indices (positions of words in table2)	Word	Count	Codes
0	0	AAAA	4	01
1	1	AATT	2	10
2	5	GATA	2	0001
3	2	ACAC	1	0010
4	3	CCGG	1	1101
5	4	CGAC	1	1110

At the end there are total five files for secure and efficient storage of fasta/multi-fasta file (T) W2, W3, W4, W5 & secret key used for encryption of W5. All these files are finally archived by "7zip" (an open source archival) to generate a single compressed output file (output.7zip) for input genome sequence T.

Step 2. Decompression Process

\\ **Input:** Output.7zip

\\ **Output:** Genome Sequence T

Step 2.1: Files W2, W3, W4, W5 & secret key are extracted from output.7zip archived file. W5 file is decrypted using the secret key to get the shuffled indices. A table "dvocab" with two columns: index and word is created. Index column is for indices of words and word column contains all the possible words (A, C, G and T) of size four stored in lexicographical order (here as an example only few entries are shown but it has 256 entries).

Step 2.2: Rearrange the contents of the "dvocab" table with the help of shuffled indices obtained by W5 file (0, 1, 5, 2, 3, and 4). A new column "new index" is also added at beginning of the Table 5.

Table 4
Contents of table "dvocab"

Index	**Word**
0	AAAA
1	AATT
2	ACAC
3	CCGG
4	CGAC
5	GATA
-	-

Table 5
Contents of table "rvocab"

New Index	Old Index	Word
0	0	AAAA
1	1	AATT
2	5	GATA
3	2	ACAC
4	3	CCGG
5	4	CGAC

Table 6
Contents of table "finaldvocab"

New Index	Old Index	Word	Codes
0	0	AAAA	01
1	1	AATT	10
2	5	GATA	0001
3	2	ACAC	0010
4	3	CCGG	1101
5	4	CGAC	1110

Step 2.3: Codes are assigned to each word of table 5 using MEncode method (algorithm1) and stored in a new column "codes". Once the "finaldvocab" table is ready decoding of binary words can be started. At any step i the decoder reads a binary word (two bits at a time) from W4 file till it ends with 01 or 10, which is decoded with the help of "finaldvocab" table (Table 6).

Compressed form of sequence T'': 01110110000101111010000101010010

Genome Sequence T': AAAACCGG AATT GATA AAAA CGACAATTGATAAAAAAAAAACAC

With the help of W3 file, positions of all 'N's (non A/C/G/T) are added in T'. W3 file contains 13, 6,27,4. To get back the original genome sequence T, non acgt character 'N' is added at positions 13 and 27 of cluster size 6 & 4 respectively. Thus the obtained sequence is:

Original genome sequence,

T= AAAACCGGAATTNNNNNNGATAAAAAN
NNNCGACAATTGATAAAAAAAAAACAC

This coding technique is a prefix free, no codeword is a proper prefix for any other code, and hence it is directly decodable. In MWBTC every code is ended with either 01 or 10, this property enables searching substrings directly over compressed data.

Encryption and Archival

Contents of W5 file are sufficient to decode (to get back the original genomic sequence) from compressed data (W4). To secure whole genome sequence it is proposed to encrypt the contents of W5 file only (Figure 5). AES 256 (Mahajan & Sachdeva, 2013) is used here to encrypt the content of W5. As the size of W5 file is very small so encryption will not take much time.

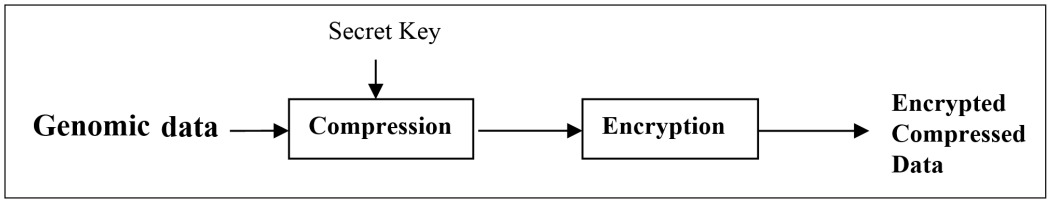


Figure 4. Existing two level approach for secure storage of genomic data: Encryption after compression

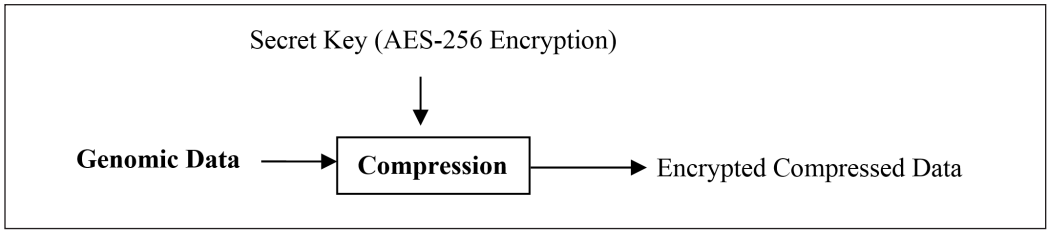


Figure 5. Proposed one level approach (unified) for secure storage of genomic data: Encryption during compression itself

In the proposed method encryption is applied only on the contents of W5 file (having indices from 0-255) only while other approaches apply the encryption on the whole compressed data (about 705Mb in case of Human genome). Thus proposed method requires very little additional computational effort to provide the security of genomic data.

After applying compression and encryption there are five files W2, W3, W4, W5 (encrypted shuffled indices) & secret key. These all are finally archived by "7zip" (an open source archival) to generate a single compressed output file (outut.7zip) for the input genome sequence.

Decompression

Files W2, W3, W4, W5 & secret key are extracted from output.7zip archived file. W5 file is decrypted using the secret key. A temporary "dvocab" table is created as in step 2.1 of example 1. Reshuffle the content of the "dvocab" table with the help of shuffled indices obtained by W5 file as shown in step 2.2 of example1. Once the "finaldvocab" table is ready decoding of binary words can begin immediately as shown in step 2.3 of example 1. From W4 file decoder reads binary words ending with 01 or 10 and decode it with the help of "finaldvocab" table.

Algorithm1

//MEncode method use to encode the binary word into binary code words

1. T is the total no of segment present in vocab
2. i is the level of n^{th} segment

```
3.  T[0] ← 01
4.  T[1] ← 10
5.  x← 0,y←0,z← 2
6.  for  i ←1 to 8
7.  for  j←0 to power(2,i)
8.  T[z].append(T[0])
10. T[z+1].append(T[1])
11.  z←z+1
12.  j←j+1
13.  i←i+1
14.  for j← 0 to power(2,i)
15.  T[k].append(T[1])
16.  T[k].append(T[y++])
17.  z←z+1
18.   j←j+1
19.  x←z-power(2,i+1)
20.  y←z-power(2,i+1)
21.  return T[i]
```

RESULT AND DISCUSSION

The performance of the proposed method has been compared with two state of the art fasta/ multi-fasta file compressors: Seqcompress and Cryfa. Four different datasets (FAN (ftp://ftp.ncbi.nih.gov/genomes/ Bacteria/all.fna.tar.gz), FEN (ftp://ftp.ncbi.nih.gov/genomes/ Bacteria/all.ffn.tar.gz), Camera (goldenPath/hg18/Chromosomes/) and Eukaryotic (https://portal.camera.calit2.net) are used for experiments. FAN, Camera and Eukaryotic datasets are in Fasta file formats while Fen dataset is in Multi-fasta file format. All experiments are performed on Ubuntu (64-bit) machine having a 2.33 GHz core i7- processor. The proposed method has been implemented in Java.

Table 7 confirms that compression ratio of proposed method is higher than other state of art algorithms. Figure 6 shows the times required for the SeqCompress, Cryfa and proposed method. SeqCompress and Cryfa are based on encryption after compression policy (Figure 4) while the proposed method uses unified compression-encryption policy (Figure 5). From Figure 6 it is confirmed that unified compression-encryption technique takes significantly less time in comparison to other techniques. Figure 7 shows memory consumption for the SeqCompress, Cryfa and proposed method. From Figure 8, it is confirmed that memory consumption of proposed method is least among existing techniques.

Table 7
Comparison of compression ratio of proposed method with SeqCom-press and Cryfa

Datasets	Uncompressed Size(in Megabits)	Compressed size using proposed method (in megabits)	Compression ratio**		
			Proposed	SeqCompress	Cryfa Method
FNA Datasets	40100.16	9155.92	**4.37**	4.21	3.89
FEN Datasets	38338.80	8168.48	**4.69**	4.34	*
Camera Data sets	19144.88	2102.24	**9.10**	8.65	7.64
Eurokaryotic	114703.68	22922.8	**5.03**	4.83	4.25

Note: *This method does not support multifasta file; ** CR=Uncompressed/Compressed

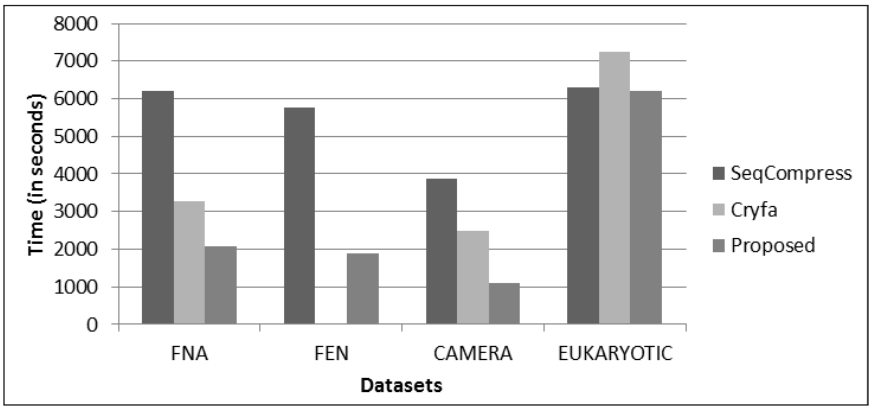


Figure 6. Time required for SeqCompress, Cryfa and Proposed method to securely and efficiently store various data sets (Time in seconds) (Cryfa does not support FEN dataset)

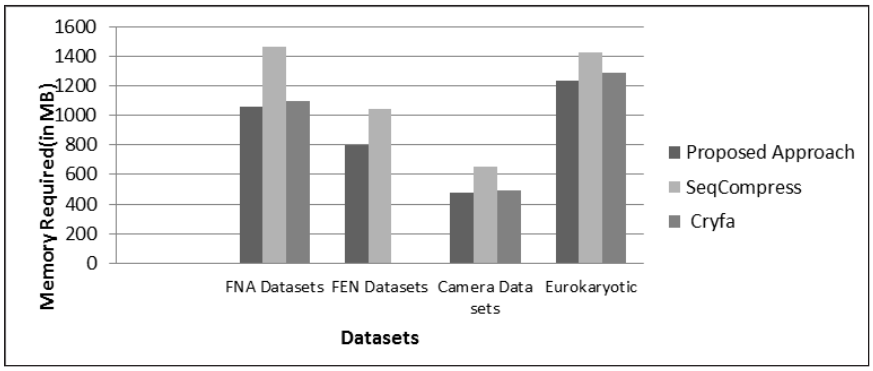


Figure 7. Memory required for SeqCompress, Cryfa and Propose method to securely and efficiently store various data sets (Memory in MB) (Cryfa does not support FEN dataset)

CONCLUSION AND FUTURE DIRECTION

In this paper, a new method to compress and encrypt the fasta/multifasta file is presented. The proposed method uses unified encryption approach (encryption during compression) to encrypt the genomic data. Here encryption is applied only on the permuted indices (0-255) created by MWBTC at the time of encoding instead of applying it on the whole genomic data. Experimental results show that proposed method has good compression ratio. Proposed algorithm also provides security of genomic data with very little computational effort. As a future work it can be extended for referential genome compression technique also.

REFERENCES

Bose, T., Mohammed, M. H., Dutta, A., & Mande, S. S. (2012). BIND–An algorithm for loss-less compression of nucleotide sequence data. *Journal of Biosciences, 37*(4), 785-789.

Chen, X., Li, M., Ma, B., & Tromp, J. (2002). DNACompress: fast and effective DNA sequence compression. *Bioinformatics, 18*(12), 1696-1698.

Danek, A., & Deorowicz, S. (2018). GTC: how to maintain huge genotype collections in a compressed form. *Bioinformatics, 34*(11), 1834-1840.

Dix, T. I., Powell, D. R., Allison, L., Yaeger, S., Bernal, J., & Stern, L. (2006). Exploring long DNA sequences by information content. In *Probabilistic Modeling and Machine Learning in Structural and Systems Biology* (pp. 97-102). Tuusula, Finland.

Dorok, S., Breß, S., Teubner, J., Läpple, H., Saake, G., & Markl, V. (2017). Efficient storage and analysis of genome data in databases. In, B. Mitschang, D. Nicklas, F. Leymann, H. Schöning, M. Herschel, J. Teubner, … & M. Wieland, (Eds.), *Datenbanksysteme für Business, Technologie und Web (BTW 2017)* (pp. 423-442). Gesellschaft für Informatik, Bonn.

Gupta, A., & Agarwal, S. (2008). A scheme that facilitates searching and partial decompression of textual documents. *International Journal of Advanced Computer Engineering, 1*(2), 99-109.

Hosseini, M., Pratas, D., & Pinho, A. J. (2016). A survey on data compression methods for biological sequences. *Information*, *7*(4), 56-78.

Jahaan, A., Ravi, D. R. T. N., & Arokiaraj, S. P. (2017). A Comparative Study and Survey on Existing DNA Compression Techniques. *International Journal of Advanced Research in Computer Science (IJARCS), 8*(3), 732-735.

Kumar, S., Agarwal, S., & Prasad, R. (2015, May). Efficient Read Alignment Using Burrows Wheeler Transform and Wavelet Tree. In *2015 Second International Conference on Advances in Computing and Communication Engineering (ICACCE)* (pp. 133-138). Dehradun, India.

Mahajan, P., & Sachdeva, A. (2013). A study of encryption algorithms AES, DES and RSA for security. *Global Journal of Computer Science and Technology, 13*(15), 14-22.

Mohammed, M. H., Dutta, A., Bose, T., Chadaram, S., & Mande, S. S. (2012). DELIMINATE—a fast and efficient method for loss-less compression of genomic sequences: sequence analysis. *Bioinformatics, 28*(19), 2527-2529.

Pinho, A. J., & Pratas, D. (2013). MFCompress: a compression tool for FASTA and multi-FASTA data. *Bioinformatics, 30*(1), 117-118.

Pratas, D., Hosseini, M., & Pinho, A. J. (2017, June). Cryfa: a tool to compact and encrypt FASTA files. In *International Conference on Practical Applications of Computational Biology & Bioinformatics* (pp. 305-312). Cham: Springer.

Sardaraz, M., Tahir, M., Ikram, A. A., & Bajwa, H. (2014). SeqCompress: An algorithm for biological sequence compression. *Genomics, 104*(4), 225-228.

Wandelt, S., Bux, M., & Leser, U. (2014). Trends in genome compression. *Current Bioinformatics, 9*(3), 315-326.

Wiese, L., Schmitt, A. O., & Gültas, M. (2018). Big Data Technologies for DNA Sequencing. In S. Sakr & A. Zomaya (Eds.), *Encyclopedia of Big Data Technology*. Cham: Springer.

Pertanika J. Sci. & Technol. 26 (4): 1927 - 1949 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Plan for the implementation of Mechanistic-Empirical Pavement Design Guide in Turkey

Saadoon Obaid Eyada[1,2]* and Osman N. Celik[2]

[1]*Anbar Health Directorate, Iraqi Ministry of Health, Ramadi, Anbar, Iraq*
[2]*Civil Engineering Department, Seljuk University, Selçuklu, Konya, 42000, Turkey*

ABSTRACT

Since 1958, most of the world's countries including Turkey depend on the Empirical Pavement Design Method established by AASHO (which is now known as AASHTO). The limitation of the data used for preparing AASHTO 1993 and Non-Mechanical Based procedure are the main reasons for the needs of new design procedure. The new AASHTO design procedure MEPDG has been established in 2002 and adopted by most of the states in the USA which is based on Mechanistic-Empirical (M-E) principles. The aim of this paper is to prepare a plan for the implementation of MEPDG in TURKEY starting with Third Region. The plan consists of two stages. Stage one is concerned with data collection and preparing of input files. This stage is subdivided into three tasks namley, Climate File, Material File and Traffic File. Stage two is associated with the use of the files prepared in Stage one to conduct studies using MEPDG software. These studies are recommended by AASHTO to evaluate the applicability of the procedure and the recommended enhancements. The studies of this stage are: (a) Sensitivity Analysis of MEPDG to Design Inputs; (b) Comparison of specific Third Region Designs with MEPDG designs; and (c) Calibration of Performance Models for Third Region in Turkey. The outcomes of this paper can be used as guidance for further studies on the implementation of MEPDG in other Regions in Turkey. Also the results of these researches can be assembled to implement the procedure for whole Turkey.

Keywords: AASHTO MEPDG, Mechanistic-Empirical, pavement design, Turkey

ARTICLE INFO

Article history:
Received: 27 January 2018
Accepted: 05 July 2018
Published: 24 October 2018

E-mail addresses:
saadooneyada@gmail.com (Saadoon Obaid Eyada)
oncelik@selcuk.edu (Osman N. Çelik)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Since 1958, most of the world countries including Turkey depend on the Empirical Pavement Design Method established by AASHO (which now known as AASHTO). During this period, the method passed through stages of enhancements in 1972, 1986, and the final set of the procedure which is still used until now is the version of 1993. In 1996, the National Cooperative Highway Research Program (NCHRP) sponsored a Project 1-37A to develop a new design guide for pavement structures. This procedure is called Mechanistic-Empirical Pavement Design Guide (MEPDG). The design guide recommended by the project team in 2004 is based on mechanistic-empirical (M-E) principles (Baus & Stires, 2010; Carvalho & Schwartz, 2006; Coree, Ceylan, & Harrington, 2005; Highway & Officials, 1993; Mallela, Quintus, & Smith, 2004).

AASHTO 1993 Design Method

The procedure is dependent on Empirical models (Equations) that relate smoothness of the pavement which can be represented by Pavement Serviceability Index PSI with Material properties, Environment, Traffic (ESAL) and other factors.

One of the limitations of AASHTO 1993 (A93) is that the data that had been used for the verification of Empirical models had been collected from limited sections in two states in the USA with low traffic loadings (less than 2 millions) and moderate environmental condition (Climate).(Carvalho & Schwartz, 2006; Highway & Officials, 1993).

MEPDG Method

The method depends on determination of responses (Strains, Deflections and Stresses) induced by passing tire load on pavement section. The Multi-Layer system or Finite Element approaches are used for the calculations of the responses. This part of the processes represents the Mechanistic part of the procedure. The Empirical part consists of the calculation of Distresses/Roughness of the pavement section due to the applied load repetitions using Transfer functions that correlate responses with Distresses/Roughness. The method can be summarized in the chart shown in Figure 1 (Baus & Stires, 2010; Carvalho & Schwartz, 2006; Papagiannakis, Bracher, Li, & Jackson, 2006).

AASHTOWare (DARWin) Software

The Mechanistic-Empirical calculations in the MEPDG cannot be performed by hand or simple spreadsheets. DARWin is a Windows-based software developed as a part of AASHTO program (AASHTOWare). This software is used for the implementation of the M-E PDG methodology. A screenshot of the software's *Main screen* is shown in Figure 2 (AASHTO, 2017; Kaya, 2015).

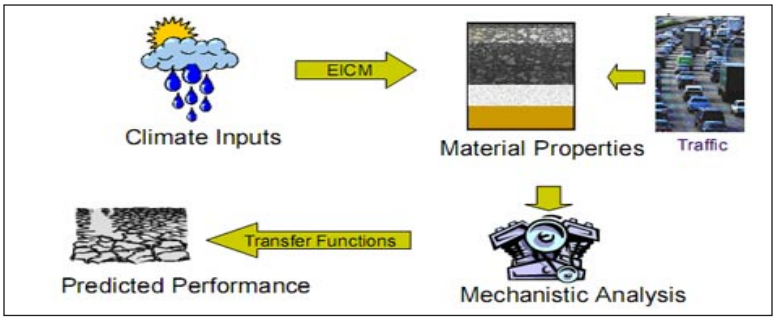


Figure 1. The Stages of MEPDG Procedure

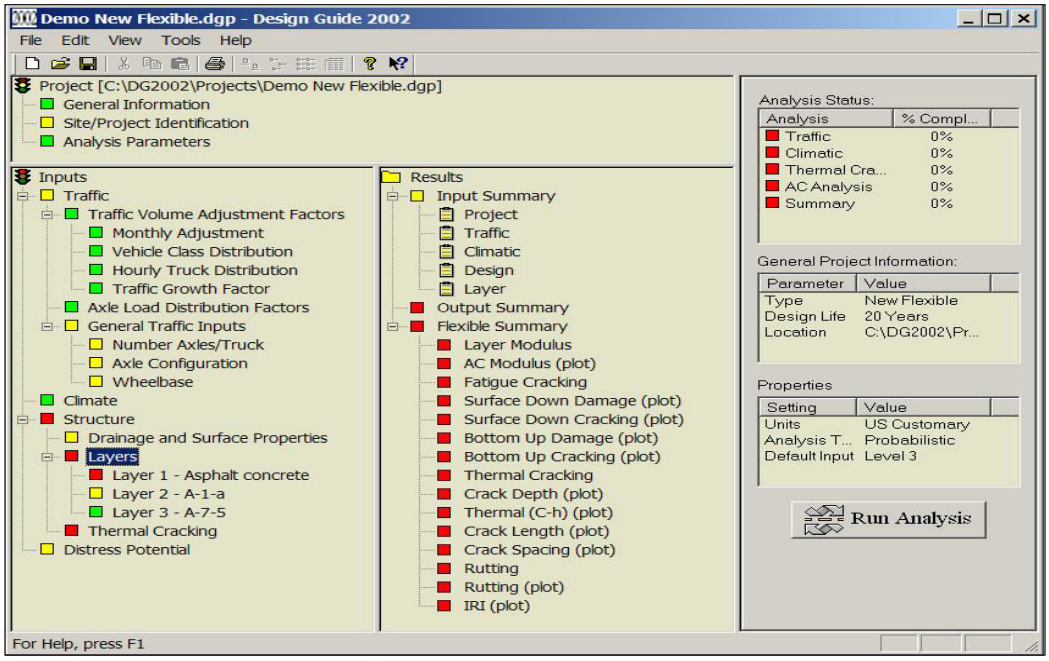


Figure 2. AASHTOWare (DARWin) Software Screenshot

Study Area – Third Region

Turkey, a bridge between two continents, is located between Europe and Asia as shown in Figure 3. In Turkey, there are 18 Regions of Transportation. One of them is the Third Region R3 as shown in Figure 4. R3 is one of the important regions. The percentage of traffic loading is between (8-10) %. The area of R3 is about 8.8 % of the total area of Turkey (http://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/Root/default.aspx).

As a starting point, R3 is selected to start the implementation of MEPDG project in Turkey. The R3 is a large area consisting of four states (Konya, Aksaray, Afyonkarahisar, and Karaman). The roads net consists of about 10% of Turkish highway roads net.

Figure 3. Geographical location of Turkey (Google Map)

Figure 4. Third region in Turkey

Geographical and Environmental Specifications

This region is in the middle of Turkey. The Geographical and Environmental specifications of region can be discussed in subsequent items.

Temperature. The temperature variation between maximum and minimum can be shown in Figures 5 and 6 (KAYA, 2011).

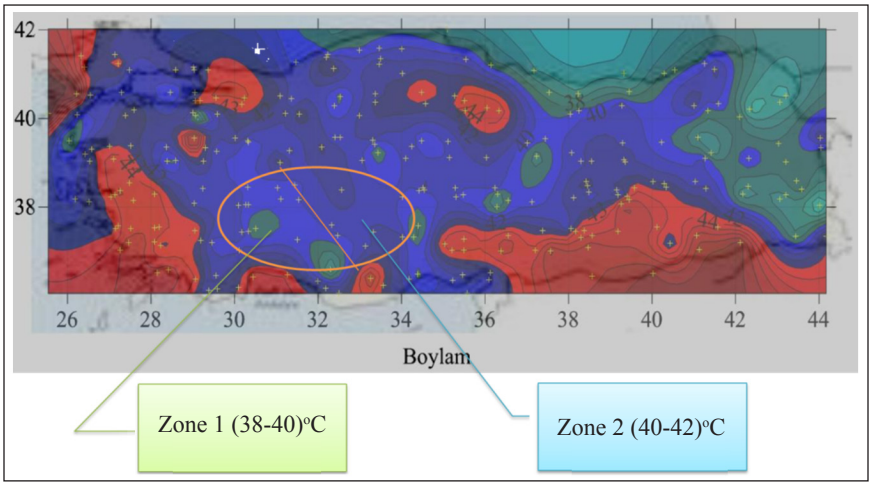


Figure 5. High temperature variation in R3 in Turkey

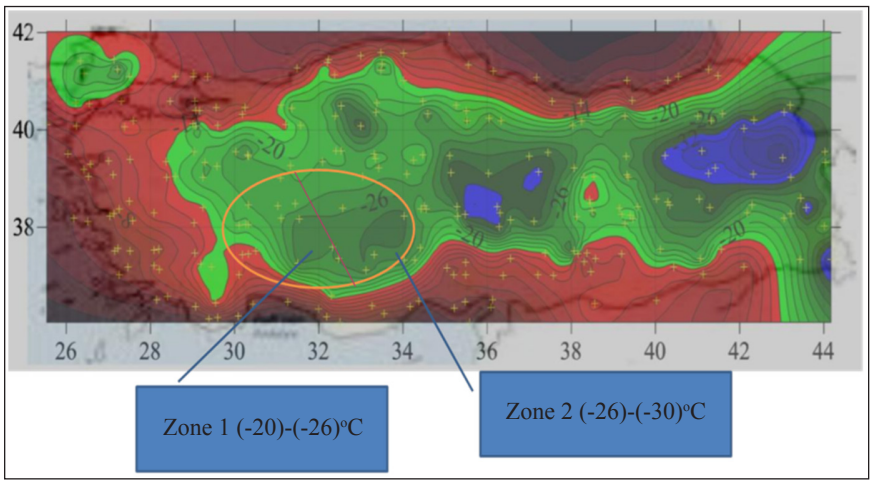


Figure 6. Low temperature variation in R3 in Turkey

As it is seen in the figures above, the weather in R3 is cold in winters when the temperature may go 30°C below zero. For that reason, the fatigue cracks are the main distress in this region.

Precipitation. The precipitation rates in R3 can be shown in Figure 7. The precipitation rates range between 300-400 mm to be the lowest rate in Turkey. The precipitation rate of R3 is lower than other regions in Turkey. This is an indicator of deep water table and low sensitivity soils to the variations in precipitation rates.

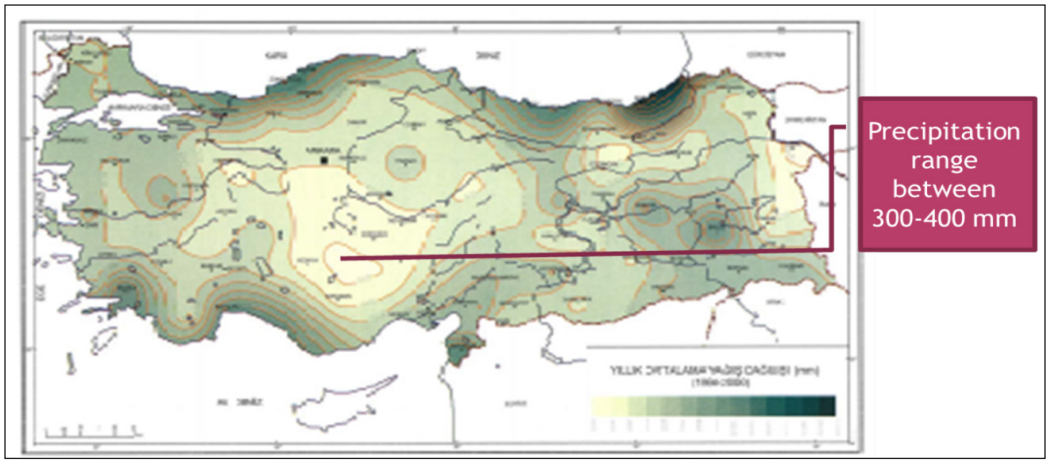


Figure 7. Precipitation variation within R3 in Turkey

Summary

1. The R3 can be environmentally classified into two regions, West and East.
2. The maximum temperature is between 38-42^{o}C, while the minimum temperature is between (-20)-(-30)^{o}C. Accordingly, asphalt type PG (56-16) is the most suitable for the R3.
3. The precipitation range of the region is the lowest in Turkey with a range of (300-400) mm.
4. The R3 is between the Latitude (30-34) and the Longitude (37-39), while the elevation is between (1000-1200) m.

Implementation Of MEPDG

The implementation of MEPDG in R3 as a first step toward the full implementation of MEPDG in Turkey must be carefully studied and divided into Stages. Each stage consists of many tasks that should be implemented simultaneously and gradually.

Each task represents a Research or a Project that either implemented by the University or by a Contractor in collaboration with Turkish Highway Administration. The Stages and their tasks can be presented in subsequent sections.(Coree et al., 2005; Flintsch, Loulizi, Diefenderfer, Diefenderfer, & Galal, 2008; Pierce, Jackson, & Mahoney, 1993). The stages and steps of the implementation are illustrated in a flow chart in Figure 8.

Stage 1

This stage is very important and the accuracy of the next stages is highly dependent on this stage. The aim of this stage is to collect the data required by MEPDG that represents the Climate, Material and Traffic of local conditions of Turkey. This data is used in

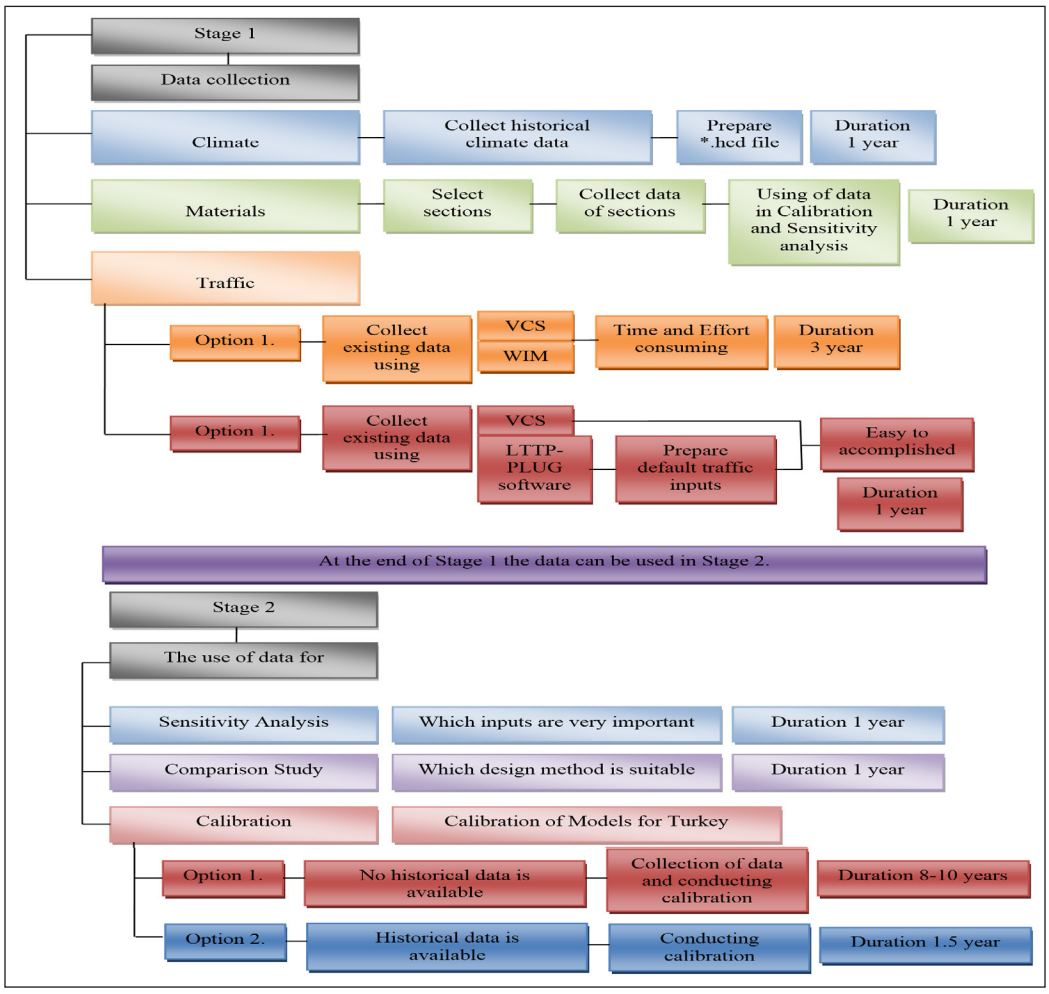


Figure 8. Implementation plan Flow Chart

subsequent steps for the implementation of MEPDG in Turkey such as Sensitivity Analysis, Comparison and Calibration. This stage consists of the following Tasks.

Task 1 (Climate File). One of the main inputs of MEPDG is the climate inputs. The Enhanced Integrated Climate Models (EICM) uses these inputs to calculate the changes in materials properties (e.g HMA modulus, Unbound material Mr) during the seasons and day. This task consists of the following steps: (Ahmed, Marukic, Zaghloul, & Vitillo, 2005; Flintsch et al., 2008; Johanneck & Khazanovich, 2010; Q. Li, Wang, & Hall, 2010; Zapata, 2009; Zapata & Houston, 2008).

Step 1: Collect climate data from weather stations around the R3 for the last five years. The data consists of Hourly (Temperature (F),Wind speed (mph)(W), % Sun shine

(SS), Precipitation (P) and Relative humidity (RH)). Below is an example about the form of data that should be entered into the MEPDG software.
1997060100,57.9,9,0,0.2,97
YYYYMMDDHH,F,W,SS,P,RH

Step 2: Completion of Un-Exist data by interpolation or averaging. Generate the complete Data file in the format of ASCII text.

Step 3: Conversion of the ASCII text file to (*.hcd) format that would be accepted by MEPDG software.

Step 4: The new added stations should be added to the Station.dat file so that the MEPDG software can read it.

Data Resource 1. The data can be obtained from the Meteorology Administration. In R3 there are 18 stations. This number of station is very good, and the virtual stations can be made easily by MEPDG software which can represent a climate condition for each site. Figure 9 shows the map of climatic stations of R3.

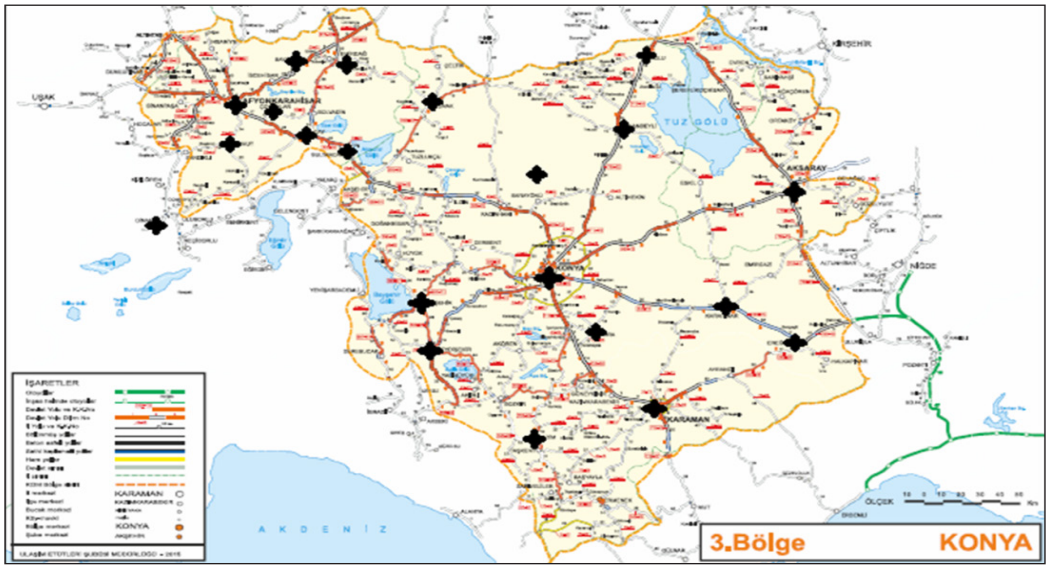


Figure 9. The weather station distribution in R3 in Turkey

Water Table 2. The climate file also consists of Water Table data. It is recommended to collect data about the water table in the R3 that can be used instead of Default values of MEPDG (Hall & Beam, 2005).

The required Water Table data is either:

1. Annual: This means that one value can represent the water table level of the site for one year (Average). Or

2. Seasonal: That means the water table is different through the seasons of the year so different values of the WT must be collected to be used by the MEPDG.

The data can be collected from: (a) laboratories; (b) universities; (c) highway administration; or any other resource of data.

Task Duration 3. This task needs about 8 months and can be established in 2018-2019.

Task Responsibility 4. It is appropriate to be carried out by University.

Task 2 (Materials File). This file is important and expensive especially if Level one is to be used. Many Departments of Transportation DOTs in the USA started the implementation of MEPDG using a mixture of Levels (2+3). Accordingly, as a first step, the Material files will be prepared according to mixture of Levels (2+3).This task consists of the following steps:(Bari, 2005; Chehab & Daniel, 2006; Flintsch et al., 2008; Hall & Beam, 2005; Mallela et al., 2004; Ovik, Birgisson, & Newcomb, 2000; Uzan, 1998).

Step 1: Select sections from the R3 to prepare the Sampling Template. These sections should be selected according to:

1. Type of Pavement (New, Overlay), or according to layer thickness (less or more than 10 cm, for example).
2. Type of roads (Interstate, Provincial).
3. Geographical Location (West, East).
4. Type of Unbound material used (HMA over Base, HMA over Base and Subbase… etc).
5. Traffic loads (High, Medium or Low).

The preparation of Sampling Template will be very useful in the Calibration process that will be conducted in the final Tasks.

It is preferable to take the following notes in consideration:

1. The number of sections can be 26 or more or according to minimum number of sections required for calibration.
2. The selected sections shall be of available data.
3. The selected sections must have different distress conditions, and it is preferable that the distress in each section is not less than 50%.

The Sampling Template can be in the form of examples shown in Tables 1 and 2.

Step 1: After the Sampling Template has been made and the sections have been selected, the data about the materials used in each section should be collected and arranged in the form required by MEPDG.

The data sources are: (a) Pavement Management System PMS; (b) Data stored in laboratories; (c) Region Three Highway Administration; (d) Any other data resource. Since the Mixture of levels (2+3) has been chosen as an implementation level, the required data is shown in Table 3. The other required properties can be used as Defaults of the MEPDG.

Table 1
The sampling template (Example1)

Thickness (cm)	**Granular base thickness (cm)**	**Subgrade type**	
		Coarse-grained soils (AASHTO Class A-1 through A-3)	**Fined-grained soils (AASHTO) Class A-4 through A-7)**
Less than 10	Less than 20		
	More than 20		
More than 10	Less than 20		
	More than 20		

Table 2
The sampling template (Example2)

Pavement type		**Structure type**	
		HMA over Base	**HMA over (Base + Subbase)**
New	Modified	3*	
	Pure		
Overlay	Modified		
	Pure		

Note: *This number indicates that there are three new sections of Modified asphalt and the pavement consists of HMA over Base.

Table 3
The sampling template of materials' properties (Example)

Material type		Required data	Level
Unbound layers (Base, Subbase) and Subgrade	Strength properties	CBR Or aggregate gradation	Level 3
	ICM properties	Aggregate gradation Atterberg limits	Level (2 + 3)
Asphalt Pavement	HMA mixture	HMA gradation %AV $\%P_{eff}$ Mix density	Level (2 + 3)
	Asphalt binder	Penetration or viscosity	Level 3

Recommendation 1. Ahmet SAĞLIK and A.Gurkan GUNGOR, conducted a study titled," RESILIENT MODULUS OF UNBOUND AND BITUMINOUS BOUND ROAD MATERIALS"(SAĞLIK and GUNGOR).The aim of their study was to measure the Resilient Modulus M_R of unbound materials and the Dynamic Modulus of HMA mixture E. They collected and tested samples of unbound materials from all the 17 regions in Turkey. The results of the study were important. They found an Empirical relationship for the determination of resilient modulus of unbound materials as shown in Eq.1 (SAĞLIK and GUNGOR).

$$M_R=175(DAC+k)0^{0.436}.CBR^{0.4}.\left[\frac{1}{1+\log(No200)}\right]^{0.35(LL.PI+1)^{0.06}}.\left[\frac{\gamma\, max}{No4}\right]^{0.09\log(\omega\, opt)} \quad (1)$$

where:
DAC: Total thickness of HMA, cm.
CBR: Soaked California Besring Ratio, %.
$\omega\ opt$: Optimum moisture content, %.
$\gamma\ \max$: Maximum Dry unit weight, g/cm^3.
LL: Liquid limit, %.
PI: Plasticity index, %.
No200: Percent passing No.200 sieve.
No4: Percent passing No.4 sieve.
K: Depth correction factor. For Base and Subbase k=0, while for Subgrade K=17

According to the same study, it was concluded that simplified Witczak's dynamic modulus equation for 4 Hz frequency may safely be used for the estimation of Dynamic modulus of Asphalt Concrete as shown in Eq.2.

$$\text{Log } E^*=3.75+0.029No200-0.00177No200^2-0.0028No4-0.058V_b-0.8\left[\frac{Vb}{Vb+Vh}\right]+\frac{3.87}{1+e^{-2.56+0.89\log(pen)-0.0015[\log(pen)]^2}}\ \underline{055No3/4} \quad (2)$$

where:
E*: Dynamic modulus, cm.
Va: Air voids, %.
V_b: Binder Content, %.
Pen: Biner penetration, 0.1 mm.
No200: Percent passing No.200 sieve.
No4: Percent passing No.4 sieve.
No3/8: Percent passing No.3/8 sieve.
No3/4: Percent passing No.3/4 sieve.

Conclusions 2. Based on the information presented earlier, there are two options:

1. The use of the results of the study by SAĞLIK and GUNGOR, so the recommended equations will be used for the determination of Mr and *E*. Or,
2. Conducting new study according to Steps 1 and 2 discussed previously to collect data and preparation of Material file (Catalogue). Accordingly the following shall be done:

Task Duration 3. This task needs about 24 months and can be established in 2018-2020.

Task Responsibility 4. It is appropriate to be carried out by joint effort between the University and Highway Administration.

Task 3 (Traffic File). The core task is how to collect traffic data and preparation of traffic file (Catalogue) for R3.The traffic data according to MEPDG is the unique of the procedure. To understand the traffic data required by the MEPDG, the following items shall be studied :(Carvalho & Schwartz, 2006; Haider, Buch, Chatti, & Brown, 2011; Harvey, Chong, & Roesler, 2000; Lu & Harvey, 2006; Lu, Zhang, & Harvey, 2009; Romanoschi, Momin, Bethu, & Bendana, 2011; Sayyady, Stone, Taylor, Jadoun, & Kim, 2010; Timm, Bower, & Turochy, 2006; Wang, Hancher, & Mahboub, 2007).

Automatic Vehicle Classification AVC Survey 1. This type of survey is used to identify:

1. Vehicle Class Distribution VCD for each type of vehicle within the traffic (e.g, Passenger cars, Trucks, Busses, etc) through the months of year.
2. Monthly Distribution Factor MDF and the distribution of each class within the year can be calculated using the results of this survey.
3. Average Annual Daily Truck Traffic AADTT.

According to FHWA classification system, there are 13 classes of vehicles as shown in Figure 10.

In Turkey there are 500 stations of VCS. The Turkish Transportation Authority has another vehicle classification system as shown in Figure 11.

The 18 Turkish classes have been grouped into five classes to simulate the FHWA classes as shown in Table 4.

It can be concluded from Table 4 that:

1. Among the 10 classes (Class4-13) according to FHWA that are used by the MEPDG, there are only 4 classes of Turkish classification system included.
2. C1 class according Turkish classification system which is opposite to Classes 1-3 according to FHWA classification system is not used by the MEPDG.

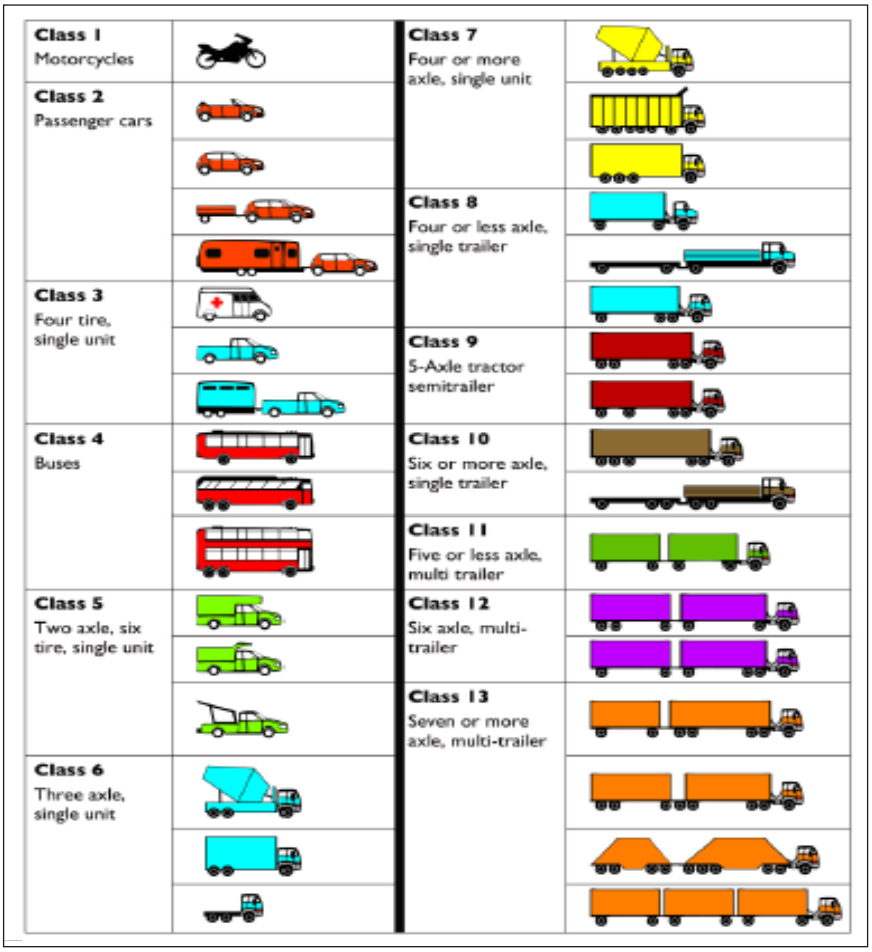


Figure 10. FHWA vehicle classification system

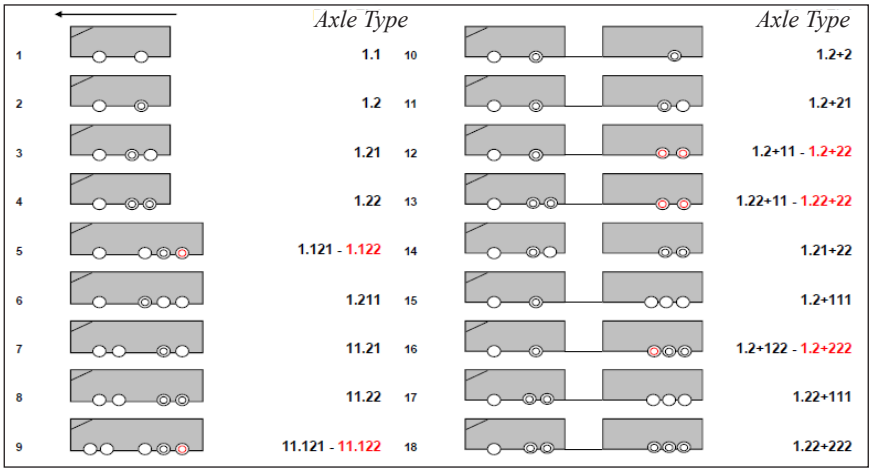


Figure 11. Turkish vehicle classification system

Table 4
Turkish as Compared to FHWA Classification System

Vehicle class as for Turkey	Vehicles included	Its classification according to FHWA
C1	Cars (less than 3.5 tons)	Class 1 + 2 + 3
C2	Medium good vehicles (3.5-10 tons)	Class 5
C3	Buss	Class 4
C4	Trucks (single unit)	Class 6 + 7
C5	Articulated (trailer)	Class 8 - 13

3. C4 class according to Turkish classification system is opposite to wide range of FHWA classes from (Class 8-13). These classes are the most important classes used by MEPDG.

A comprehensive study must be conducted in Turkey for the identification of each class of traffic according to FHWA classification system to be used by MEPDG.

Weight in Motion WIM Survey 2. This type of survey is the main difference between the A93 and the MEPDG. In this type of survey, the weight of each type of axle (Single, Tandem..) of each type of vehicle (Bus, Trailer, ..) is measured when the vehicle passes over the WIM instrument which is installed beneath the road section as shown in Figure 12 (Ramachandran, Taylor, Stone, & Sajjadi, 2011; Sayyady et al., 2010).

The most important results of this survey is the Load Spectra LS which represents the variation of load of each axle (Single, Tandem) of each Vehicle class (Class 4-13), during different seasons. Figure 13 shows an example of the results of WIM survey and load spectra concept.

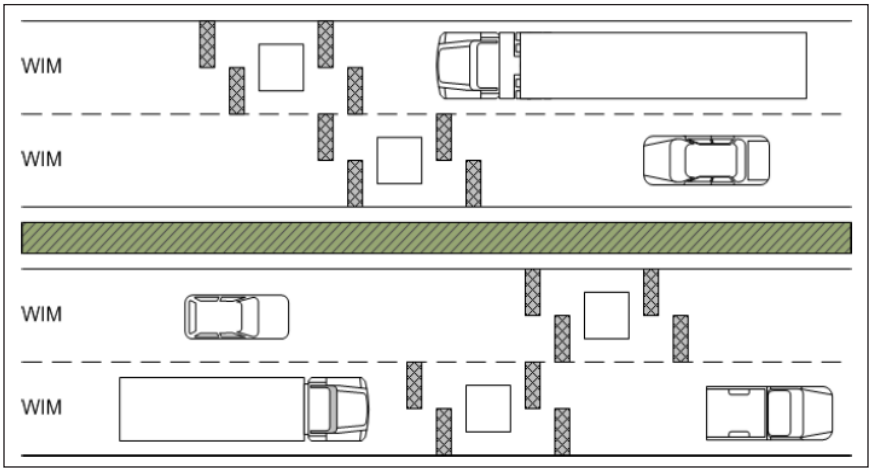


Figure 12. WIM system setup

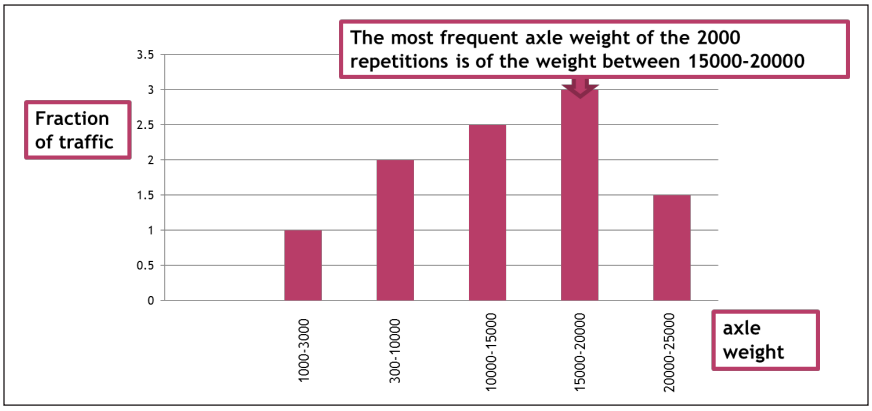


Figure 13. Example of the results of WIM survey and load spectra *Concept*

Options of Preparation of Traffic File 3. There are two options to prepare the Traffic file according to the Effort, Time and Cost available for each option.

Option 1 2a. In this option a Mixture of Level2+3 (Regional level) is selected, therefore a comprehensive study is conducted according to the following steps:

Step 1: Select sections in the same procedure mentioned in Material file preparation section to represent R3 roads. The selected sections must have permanent Automated Vehicle Classification AVC stations to complete the data required by MEPDG.

Step 2: Installing new WIM stations to collect the data required by the MEPDG according to "Traffic Monitoring Guide TMG." The TMG is the most important reference published by FHWA for this purpose(FHWA, 2013).

Step 3: The use of collected data to determine the R3 traffic default values instead of that provided by the MEPDG Level3.

The "Long-Term Pavement Performance Pavement Loading User Guide (LTPP PLUG)" is the main reference for the preparation of New Local Traffic Defaults that can be used by MEPDG to represent the traffic conditions of R3. For this option the LTPP PLUG SOFTWARE can be used for the preparation of New Local Traffic Defaults data for R3 (Selezneva & Hallenbeck, 2013).

Option 2 2b. In this option, Level3 can be selected; therefore, the default values provided by the MEPDG can be used. The most important disadvantage of this option is that the default values of level3 is based on data not representing the actual traffic loadings of R3.Accordingly, it is recommended that these defaults be changed by making a comparison between the traffic data of R3 and provided by The "Long-Term Pavement Performance

Pavement Loading User Guide (LTPP PLUG)". Based on the comparison results, a suitable New Local Traffic Default values can be selected from LTPP PLUG to represent R3 Traffic loadings. For this option the LTPP PLUG SOFTWARE can also be used.

Recommendations 4. The following recommendations are based on statistical data provided by the THPDM website.

1. The vehicle classes that frequently represent the traffic loadings in Turkey are shown in Figure 14. The classes in figure are based on FHWA system.

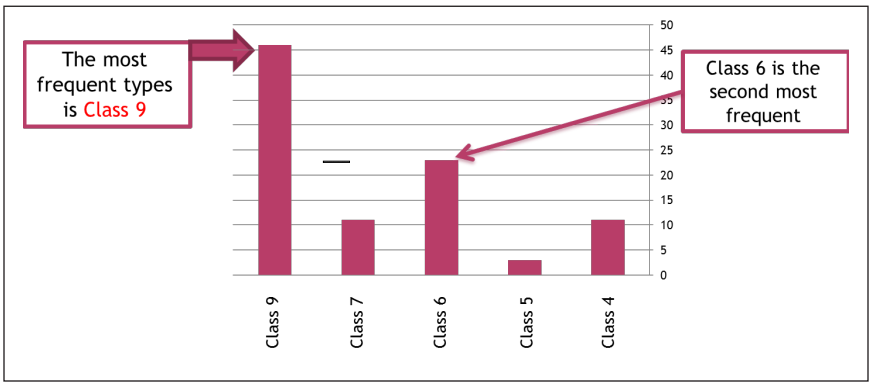


Figure 14. The most frequent vehicle classes in Turkey according to FHWA *Classes*

Figure 13, shows that there are 4 main classes to be taken into consideration in the stage of Traffic file preparation:

-Class 9
-Class 6
-Class 7
-Class 4

Concentration must be on Class 9 and 7 because they are the heaviest

2. It is important to use the available data in Turkish Transportation Administration (TTA) especially that provides data about the VCD system.
3. There are 106 stations in Turkey to measure the Average Vehicle Load. The data of these stations may be useful if there is a similarity with data that can be calculated by WIM stations.
4. The general data required by MEPDG such as (Growth factor, No Axle per Class ..) can be obtained from data available at TTA' website, and there is no need to use default values of the MEPDG.
5. The followings references and software are the most important ones that can be used for the preparation of Traffic file:

(a) Traffic Monitoring Guide TMG, from FHWA. Used for WIM survey.
(b) Long-Term Pavement Performance Pavement Loading User Guide (LTPP PLUG), from FHWA. Used for New Local Defaults of Traffic for MEPDG.
(c) LTPP PLUG SOFTWARE, from FHWA. Used for New Local Defaults of Traffic for MEPDG.
(d) TrafLoad SOFTWARE., from NCHRP, used to estimate vehicle classification statistics, load spectra, and traffic growth rates (Systematics, Center, & Corporation, 2005).
(e) MEPDG Part2 Design Inputs, Chapter4 Traffic, NCHRP Report 1-37A.

Conclusion 5. According to the selected option; either Option1 or Option2 as mentioned earlier, the followings shall be specified:

Task Duration (Option 2) 4a. This task needs about 12 months and can be established in 2018-2019.

Task Responsibility (Option 2) 4b. It is appropriate to be carried out by the University.

Task Duration (Option 1) 4c. This task needs about 36 months and can be established in 2018-2021.

Task Responsibility (Option 1) 4d. It's appropriate to be carried out by a Contractor. The QA and QC are the responsibility of TTA. The QA and QC shall be conducted according to The American Society for Testing and Materials ASTM Test Procedures.

Stage 2

After the preparation of Climate, Materials and Traffic data files, it is important to use these files to complete the Implementation processes of MEPDG. This stage consists of many Tasks that can be implemented step by step as explained in subsequent sections.

Task 1 (Sensitivity Analysis of MEPDG to Design Inputs). Not all Inputs are at the same weight of importance; some are very important and have a significant effect on the design results while others have minor effect on the results (Guclu & Ceylan, 2005; Guclu, Ceylan, Gopalakrishnan, & Kim, 2009; Hall & Beam, 2005; Kim, Ceylan, & Heitzman, 2005; McDonald & Madanat, 2011; Orobio, 2010).

Objective 1.

1. Study the effect of various Climate, Materials, Traffic, Structure, Reliability level and Calibration Coefficients inputs on the design results using MEPDG.
2. The use of the study results to focus on the inputs of high degree of importance.

Task Duration 2. This task needs about 12 months and can be established in the 2018-2019.

Task Responsibility 3. It is appropriate to be carried out by the University.

Task2 (Comparison of specific R3's Designs with MEPDG's Designs). There could be a differences between the design using MEPDG and using the design procedure followed by R3 administration (i.e. A93). Typical (3-6) sections will be chosen to represent the relevant designs in R3. A comparison between the as built designs of these sections with that of MEPDG can be made.

Objective 1. Study the differences between the MEPDG and design procedure followed by R3 Transportation Authority.

The use of the study results to recommend which procedure is more suitable and can satisfy the conditions of R3.

Task Duration 2. This task needs about 12 months and can be established in the years 2018-2019.

Task Responsibility 3. It is appropriate to be carried out by the University.

Task3 (Calibration of performance models for R3 in Turkey). This task is very important. The calibration of performance models needs a historical data about the International Roughness Index IRI and Distresses (Rutting, Cracks types) for at least three observations within the last 10 years so that the relation between the Time and IRI/Distress can be plotted and statistical solution can be achieved (Ali, Tayabji, & La Torre, 1998; Hall, Xiao, & Wang, 2011; Hoegh, Khazanovich, & Jense, 2010; F. Jadoun & Kim, 2012; Jadoun, 2011; Li, Pierce, & Uhlmeyer, 2009; Muthadi & Kim, 2008).

Objective 1. Study the needs of MEPDG performance models for local calibration to be suitable for R3 conditions.

Calibration of performance models to represent the condition of R3 in Turkey. By this way, the MEPDG can be used for the design of New and Overlay Pavements of R3 in Turkey without bias.

Calibration Options 2. The calibration (if needed) needs a historical data, so the calibration depends on the availability of historical data. Accordingly, there are two options:

Option 1: NO historical data is available 1a. This is the worst case scenario where the data is not available so that new study must be conducted to collect data of IRI/Distress. This study is a part of Pavement Management System PMS, so the study has two objectives:

1. Collecting data for Calibration of Performance Models.
2. Starting a program of PMS in R3.

The new study consists of the followings:

1. Selecting of minimum number of sections as required by MEPDG.
2. The selected sections must represent different Environmental, Road types, Traffic loads and Geographical conditions of R3 in Turkey.
3. Conducting a data collection study for IRI/Distress for these sections using Automated Pavement Distress Collection Techniques. The IRI/Distress units must be according to Long Term Pavement Performance LTPP requirements.
4. The data must be collected at least every two years within 8-10 years.
5. At the end of the study the calibration process can take place as in Option2.

Task Duration 1a-1. This task needs about 8-10 years and can be established in 2018-2026.

Task Responsibility 1a-2. It is appropriate to be carried a Contractor. The TTA is responsible for QA and QC.

Important References 1a-3.

1. NCHRP SYNTHESIS 334, "Automated Pavement Distress Collection Techniques"(McGhee, 2004).
2. LTPP pavement condition evaluation.

Option 2: Historical Data is Available 1b. This is the simplest case were a few effort and short period of time required.

The followings are the most important processes that must take place:

1. Collecting the available data from: (a) PMS source; (b) TTA; and (c) Researches and Academic institutions.
2. Arrangement of data according to LTPP units and MEPDG requirements.
3. Using of prepared data to conduct Calibration process according to "Guide for the Local Calibration of the Mechanistic-Empirical Pavement Design Guide" published by AASHTO-2010.

Task Duration 1b-1. This task needs about 18 months and can be established at the year 2019-2021.

Task Responsibility 1b-2. It is appropriate to be carried by the University.

CONCLUSION

The conclusions will be based on the results of the implementation of the previously mentioned plan. It is expected that the results will be important and have a significant effect on the transition of Turkey towards the new Mechanistic Empirical Pavement Design.

REFERENCES

AASHTO. (2017). *DARWin 2017*. Retrieved November 12, 2017, from http://www.aashtoware.org/Pavement/Pages/default.aspx

Ahmed, Z., Marukic, I., Zaghloul, S., & Vitillo, N. (2005). Validation of enhanced integrated climatic model predictions with New Jersey seasonal monitoring data. *Transportation Research Record: Journal of the Transportation Research Board,* (1913), 148-161.

Ali, H., Tayabji, S., & La Torre, F. (1998). Calibration of mechanistic-empirical rutting model for in-service pavements. *Transportation Research Record: Journal of the Transportation Research Board,* (1629), 159-168.

Bari, J. (2005). *Development of a new revised version of the Witczak E* predictive models for hot mix asphalt mixtures.* Doctoral dissertation. Arizona State University Tempe, AZ.

Baus, R., & Stires, N. (2010). *Mechanistic-Empirical Pavement Design Guide Implementation* (No. FHWA-SC-10-01). Department of Civil & Environmental Engineering, University of South Carolina.

Carvalho, R., & Schwartz, C. (2006). Comparisons of flexible pavement designs: AASHTO empirical versus NCHRP Project 1-37A mechanistic-empirical. *Transportation Research Record: Journal of the Transportation Research Board,* (1947), 167-174.

Chehab, G., & Daniel, J. (2006). Evaluating recycled asphalt pavement mixtures with mechanistic-empirical pavement design guide level 3 analysis. *Transportation Research Record: Journal of the Transportation Research Board,* (1962), 90-100.

Coree, B., Ceylan, H., & Harrington, D. (2005). Implementing the Mechanistic-Empirical Pavement Design Guide: Implementation Plan. *InTrans Project Reports*. Iowa State University.

Federal Highway Administration [FHWA]. (2013). *Traffic Monitoring Guide TMG.* New Jersey: FHWA.

Flintsch, G., Loulizi, A., Diefenderfer, S., Diefenderfer, B., & Galal, K. (2008). Asphalt material characterization in support of mechanistic-empirical pavement design guide implementation in Virginia. *Transportation Research Record: Journal of the Transportation Research Board,* (2057), 114-125.

Guclu, A., & Ceylan, H. (2005). *Sensitivity Analysis of Rigid Pavement Systems Using Mechanistic-Empirical Pavement Design Guide.* Paper presented at the Proceedings of the 2005 Mid-Continent Transportation Research Symposium, Ames, Iowa.

Guclu, A., Ceylan, H., Gopalakrishnan, K., & Kim, S. (2009). Sensitivity analysis of rigid pavement systems using the mechanistic-empirical design guide software. *Journal of Transportation Engineering, 135*(8), 555-562.

Haider, S., Buch, N., Chatti, K., & Brown, J. (2011). Development of traffic inputs for mechanistic-empirical pavement design guide in Michigan. *Transportation Research Record: Journal of the Transportation Research Board,* (2256), 179-190.

Hall, K., & Beam, S. (2005). Estimating the sensitivity of design input variables for rigid pavement analysis with a mechanistic-empirical design guide. *Transportation Research Record: Journal of the Transportation Research Board,* (1919), 65-73.

Hall, K., Xiao, D., & Wang, K. (2011). Calibration of the mechanistic-empirical pavement design guide for flexible pavement design in Arkansas. *Transportation Research Record: Journal of the Transportation Research Board,* (2226), 135-141.

Harvey, J., Chong, A., & Roesler, J. (2000). Climate regions for mechanistic-empirical pavement design in California and expected effects on performance. CAL/APT Program, Pavement Research Center, Institute of Transportation Studies, University of California, Berkeley.

American Association of State Highway and Transportation Officials (1993). *AASHTO Guide for Design of Pavement Structures* (Vol. 1). Washington, D. C.: AASHTO & Transportation Officials.

Hoegh, K., Khazanovich, L., & Jense, M. (2010). Local calibration of Mechanistic-Empirical pavement design guide rutting model: Minnesota road research project test sections. *Transportation Research Record: Journal of the Transportation Research Board,* (2180), 130-141.

Jadoun, F., & Kim, Y. (2012). Calibrating Mechanistic-Empirical Pavement Design Guide for North Carolina: Genetic Algorithm and Generalized Reduced Gradient Optimization Methods. *Transportation Research Record: Journal of the Transportation Research Board,* (2305), 131-140.

Jadoun, F. M. (2011). *Calibration of the flexible pavement distress prediction models in the mechanistic empirical pavement design guide (MEPDG) for North Carolina*. North Carolina State University.

Johanneck, L., & Khazanovich, L. (2010). Comprehensive evaluation of effect of climate in mechanistic-empirical pavement design guide predictions. *Transportation Research Record: Journal of the Transportation Research Board,* (2170), 45-55.

Kaya, M. (2011). *Determination of Performance Grade of Asphalt Binders for Turkey Geographical Regions and Consideration of SBS Surveys* (Doctoral dissertation). Selcuk University, Turkey.

Kaya, O. (2015). *Investigation of AASHTOWare Pavement ME Design/Darwin-ME TM performance prediction models for Iowa pavement analysis and design* (Graduate Thesis & Dissertation). Iowa State University, USA.

Kim, S., Ceylan, H., & Heitzman, M. (2005). *Sensitivity study of design input parameters for two flexible pavement systems using the mechanistic-empirical pavement design guide.* Paper presented at the Proceedings of the 2005 Mid-Continent Transportation Research Symposium, Ames, Iowa.

Li, J., Pierce, L., & Uhlmeyer, J. (2009). Calibration of flexible pavement in mechanistic-empirical pavement design guide for Washington State. *Transportation Research Record: Journal of the Transportation Research Board,* (2095), 73-83.

Li, Q., Wang, K. C., & Hall, K. D. (2010). Verification of virtual climatic data in MEPDG using the LTPP database. *International Journal of Pavement Research and Technology, 3*(1), 10-15.

Lu, Q., & Harvey, J. (2006). Characterization of truck traffic in California for mechanistic-empirical design. *Transportation Research Record: Journal of the Transportation Research Board,* (1945), 61-72.

Lu, Q., Zhang, Y., & Harvey, J. (2009). Estimation of truck traffic inputs for mechanistic-empirical pavement design in California. *Transportation Research Record: Journal of the Transportation Research Board,* (2095), 62-72.

Mallela, J., Quintus, H. V., & Smith, K. L. (2004). *Consideration of lime-stabilized layers in mechanistic-empirical pavement design*. The National Lime Association, Arlington, Virginia, USA.

McDonald, M., & Madanat, S. (2011). Life-cycle cost minimization and sensitivity analysis for mechanistic-empirical pavement design. *Journal of Transportation Engineering, 138*(6), 706-713.

McGhee, K. H. (2004). *Automated pavement distress collection techniques* (Vol. 334). Washington D. C.: Transportation Research Board.

Muthadi, N., & Kim, Y. (2008). Local calibration of mechanistic-empirical pavement design guide for flexible pavement design. *Transportation Research Record: Journal of the Transportation Research Board,* (2087), 131-141.

Orobio, A. (2010). *Sensitivity analysis of flexible pavement performance parameters in the mechanistic-empirical design guide*. Doctoral Dissertation. West Virginia University.

Ovik, J. M., Birgisson, B., & Newcomb, D. (2000). *Characterizing seasonal variations in pavement material properties for use in a mechanistic-empirical design procedure*. Minnesota: Department of Transportation Office of Research Services.

Papagiannakis, A., Bracher, M., Li, J., & Jackson, N. (2006). Sensitivity of NCHRP 1-37A pavement design to traffic input. *Transportation Research Record: Journal of the Transportation Research Board,* (1945), 49-55.

Pierce, L. M., Jackson, N. C., & Mahoney, J. P. (1993). Development and Implementation of a Mechanistic, Empirically Based Overlay Design Procedure for Flexible Pavements. In *Transportation Research Board* (pp. 120-128). National Research Council, Washington, D. C.

Ramachandran, A. N., Taylor, K. L., Stone, J. R., & Sajjadi, S. S. (2011). NCDOT quality control methods for weigh-in-motion data. *Public Works Management & Policy, 16*(1), 3-19.

Romanoschi, S., Momin, S., Bethu, S., & Bendana, L. (2011). Development of traffic inputs for new mechanistic-empirical pavement design guide: Case study. *Transportation Research Record: Journal of the Transportation Research Board,* (2256), 142-150.

Sağlik, A., & Gungor, A. G. (2012). Resilient Modulus of Unbound and Bituminous Bound Road Materials. In *5th Eurasphalt & Eurobitume Congress* (pp. 455-463). Istanbul.

Sayyady, F., Stone, J., Taylor, K., Jadoun, F., & Kim, Y. (2010). Clustering analysis to characterize mechanistic-empirical pavement design guide traffic data in North Carolina. *Transportation Research Record: Journal of the Transportation Research Board,* (2160), 118-127.

Selezneva, O. I., & Hallenbeck, M. (2013). *Long-term pavement performance pavement loading user guide (LTPP PLUG)*. Federal Highway Administration, FHWA-HRT-13-089, Washington D. C.

Systematics, C., Center, W. S. T., & Corporation, C. S. (2005). *Traffic data collection, analysis, and forecasting for mechanistic pavement design* (Vol. 538). Washington D. C.: Transportation Research Board.

Timm, D., Bower, J., & Turochy, R. (2006). Effect of load spectra on mechanistic-empirical flexible pavement design. *Transportation Research Record: Journal of the Transportation Research Board,* (1947), 146-154.

Uzan, J. (1998). Characterization of clayey subgrade materials for mechanistic design of flexible pavements. *Transportation Research Record: Journal of the Transportation Research Board,* (1629), 189-196.

Wang, Y., Hancher, D. E., & Mahboub, K. (2007). Axle Load Distribution for Mechanistic–Empirical Pavement Design. *Journal of Transportation Engineering, 133*(8), 469-479.

Zapata, C. E. (2009). *Consideration of Climate in New AASHTO Mechanistic-Empirical Pavement Design Guide.* Paper presented at the Transportation Research Board 88th Annual Meeting, Washington, D. C.

Zapata, C. E., & Houston, W. N. (2008). *Calibration and validation of the enhanced integrated climatic model for pavement design* (Vol. 602). Washington D. C.: Transportation Research Board.

Pertanika J. Sci. & Technol. 26 (4): 1951 - 1974 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

An Enhanced Artificial Bee Colony Based EELB-PWDGR for Optimized Route Selection in Wireless Multimedia Sensor Networks

Hasib Daowd Esmail Al-Ariki[1]*, Mohammed Abdulhalim Alareqi[2] and M. N. Shanmukha Swamy[3]

[1]JSS Research Foundation, Sri Jayachamarajendra College of Engineering, Mysore, Karnataka-570006, India
[2]Sana'a Community College, Sana'a, Republic of Yemen.
[3]Department of Electronics and Communication, Sri Jayachamarajendra College of Engineering, Mysore, Karnataka-570006, India

ABSTRACT

Energy efficient and load-balanced routing based on QoS requirements in Wireless Multimedia Sensor Network have been achieved using Energy Enhanced Load Balancing Pairwise Directional Geographical Routing (EELB-PWDGR). However, the process of path selection is a time-consuming one. Hence, Artificial Bee Colony (ABC) is used to select the optimal path satisfying QoS constraints. Though the ABC-EELB-PWDGR outperforms EELB-PWDGR, the ABC algorithm exhibits unsatisfactory performance with a lower search speed, poor population diversity, stagnation within working methods, and struck to the local optimal solution. This study uses an enhanced ABC algorithm in which the global best solution information is added into the solution search equations in order to find the new solution only around the best solutions of the previous iterations for improved exploitation. Since Population initialization is imperative due to its impact on convergence speed, an initialization approach which based on chaotic systems and opposition based learning method has been employed to balance the diversity and convergence capability of ABC. Thus the Enhanced ABC based EELB-PWDGR (EABC- EELB-PWDGR) constructs an initial population with a maximum diversity to provide the best search solutions and a high degree of accuracy. The experimental results prove that the proposed EABC-EELB-PWDGR provides better routing performance than ABC-EELB-PWDGR.

Keywords: Artificial Bee Colony, chaotic search method, EELB-PWDGR, learning mechanism, optimization, opposition-based, routing, wireless multimedia sensor networks

ARTICLE INFO

Article history:
Received: 27 January 2018
Accepted: 5 July 2018
Published: 24 October 2018

E-mail addresses:
hasibalariki@gmail.com (Hasib Daowd Esmail Al-Ariki)
alareqi_mohammed@yahoo.com (Mohammed Abdulhalim alareqi)
mnsjce@gmail.com (M. N. Shanmukha Swamy)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Wireless Multimedia Sensor Networks (WMSN) focuses on aggregating multimedia data over distant geographical locations. With efficient routing techniques, these enable networks multi-dimensional applications in tracking, home automation, and environment monitoring. Different routing approaches such as GPSR, DGR, and TPGF. have been evaluated (Al-Ariki, & Swamy, 2017) with the PWDGR (Wang, Zhang, Wang, Ma, & Chen, 2015) emerging as the best routing technique. However, the issues of high energy consumption and inefficient load balancing continue to remain. Generally, in Wireless Sensor Networks (WSN), compression techniques are used for high energy consumption problems. However, in WMSN, the maximal image compression schemes consume more energy. Additionally, the compression schemes are utilized in the camera-equipped node, making it further worse. Similarly, since the selection of nodes is mostly in trajectory lines, the general load balancing approaches cannot improve load balancing in WMSN. This results in load imbalance and also hinders multimedia transmission. EE-PWDGR and EELB-PWDGR (Al-Ariki, Mohammed, & Swamy, 2017a; 2017b) has been developed to overcome the limitations of PWDGR. EELB-PWDGR utilizing the cluster based Self Positioning Algorithms (SPA) (Srie Vidhya Janani, & Ganesh Kumar, 2015) and GPS free localization algorithms (Wang & Xu, 2010) for energy efficient localization algorithm to locate the local coordinates of the nodes in the geographically distant places. The triangulation method is then employed after merging the coordinates to locate the node positions. Following this, the loop-free paths with effective forwarding nodes based on the energy drain and distance measures are selected. The paths are then analysed for QoS values and based on these parameters, the priority is assigned. The high prioritized image contents are then allocated to the priority paths. However, EELB-PWDGR lacks the selection of an optimal route for effective transmission, which limits its overall performance.

This study utilizes the Artificial Bee Colony (ABC) algorithm (Karaboga, & Akay, 2009; Zhang, & Liu, 2015) in EELB-PWDGR to enable optimal selection of routing paths. However, the insufficiency in solution search operation of ABC reduces the efficiency. Hence, EABC-EELB-PWDGR has been introduced with two new solution search equations to reduce time latency. Additionally, the chaotic search method and the opposition-based learning mechanism have been employed in population initialization to enhance global convergence. Based on the two new solution search equations, the optimal path is chosen and the paths are compared with for ranking purposes, based on the priority within lesser search time.

Related works

Le Dinh and Nguyen (2010) proposed a greedy geographical routing with the path optimization for WSN. The approach provides optimal routing with void avoidance

capacity and also improves the lifetime using ant routing optimization, as mentioned by Hu and Zhang (2010). Paschalidis and Li (2011) proposed energy optimized topologies for providing distributed routing in WSN using the graph-based algorithms of bi-directional spanning trees. Ye and Mohamadian (2014) proposed an adaptive clustering based dynamic routing using ant colony optimization for WSN. Zheng and Luo (2014) also proposed a novel routing protocol by using ABC to address the delay-energy trade-off problem. However, the average packet delivery ratio of this approach was found to be quite less. Ding, Tang and Ji (2016) proposed an optimized routing protocol with queuing model to reduce both energy consumption and congestion in WSN. Chatterjee and Das (2015) proposed Ant Colony Optimization (ACO) based dynamic source routing for MANETs in order to provide high packet delivery ratio with reduced delay. Selvi and Latha (2015) developed lifespan conservation method for the heterogeneous WSN using the ABC algorithm. However, in this method, the computational cost was found to increase due to the population of solutions. Cai, Duan, He, Yang and Li (2015) proposed Bee-Sensor-C, an energy-aware and scalable multipath routing protocol for wireless sensor networks based on the dynamic cluster and foraging behaviour of the bee swarm in order to balance energy consumption.

Azharuddin and Jana (2016) proposed the utilization of Particle swarm optimization (PSO) for the effective maximization of the network lifetime in WSN. This approach includes both routing and clustering algorithms to ensure energy efficient selection of cluster head and effective load balancing. Similarly, Xiang, Wang and Zhou (2016) proposed an energy-efficient routing algorithm for software-defined wireless sensor networks using the PSO algorithm. However, PSO easily suffers from the partial optimism, which results in lesser speed and direction for particles. Ari, Yenke, Labraoui, Damakoa and Gueroui (2016) introduced the concept of honey bee swarm intelligence based approach using ABC for achieving power efficient cluster-based routing. Zeng and Dong (2016) introduced an improved harmony search concept for the energy efficient routing in WSN, which also enhanced both the convergence speed and the accuracy of routing. However, since the vital QoS parameters were not well addressed, this approach seems directionless.

Rao and Banka (2017) had developed a routing algorithm based on chemical reaction optimization which clustered the network and optimized routing paths. However, this model does not consider the fault-tolerance and delay for forwarding the data which is a major drawback. Sirdeshpande and Udupi (2017) developed fractional lion optimization for cluster head based energy efficient routing. Similarly, Sohan, Mittal, Singh and Sohi (2018) also proposed a tree based routing model using PSO but this model had issues in load balancing. To their account, Gupta and Jha (2018) developed an integrated model with clustering using improved cuckoo search and routing using improved harmony search. However this sustainability of this model with communication voids and delay sensitivity

was not examined. This literature survey provides an overview of the advantages and disadvantages of the optimization load balancing-based routing methodologies, whose functional scenarios are analysed for developing the approaches proposed in this study.

MATERIALS AND METHODS

Since EELB-PWDGR approach consumes more time to compare the routing paths based on the QoS values, waiting time for the high priority multimedia data transmission tends to remain considerably high. In order to overcome this problem, the Optimization algorithms have been utilized in this study. The Artificial Bee Colony based EELB-PWDGR (ABC-EELB-PWDGR) has been proposed to reducing the time consumed for route selection. ABC-EELB-PWDGR has similar processing stages as in EELB-PWDGR, the major difference being the route selection mechanism. However, ABC has certain drawbacks of ABC, which have been overcome in EABC-EELB-PWDGR.

Optimization-based EELB-PWDGR

Initialization. In the initialization process, the early assumptions are made for localising the nodes in the network without involving the use of the GPS approach. The following assumptions are made to detect the node locations and form efficient routes.

The observed network is a network of wireless devices
All the nodes are stationary
There are no landmarks for absolute location information of a node.
All nodes have the same technical characteristics
All wireless links between the nodes are bidirectional
The nodes use Omni-directional antennae
Priority is assigned to the multimedia data packets

The improved SPA such as cluster-based SPA and Backbone-based SPA (BSPA) have been employed for finding the geographical coordinates of the sensor nodes. The sensor nodes are deployed in a random manner; the node specialization as master, slave, and border nodes is performed using a waiting timer for each node. Following this, the image or video data packets to be transmitted are assigned as high, low and medium priority depending upon the quality and impact of the particular packets in the overall aggregation performance. These prioritized data packets are transmitted over the priority routing paths at the end of the route selection stage.

Energy Efficient Node Localization

The energy efficient node localization algorithm (Wang & Xu, 2010) focuses on detecting the local coordinates of the nodes and using them determines the geographical positions

of the nodes. In the first stage, each sensor node of the network initiates its own local coordinate system based on the triangulation method. It then determines the coordinates of other nodes with respect to its coordinate system. Thus, the local coordinate systems of each node are determined. In the second stage, the individual coordinate systems are merged to form the global coordinate system. The transformation equation consisting of the transformation matrix is computed for each global coordinate system, whose solution is determined by the direction adjustment. The positions of each node can be located by resolving the transformation equations.

Route Construction

Detection of node locations is followed by route construction. The paths are constructed considering loop-free and disjoint route conditions. In order to ensure a loop-free path, the invariant conditions based on the sequence number, distance to destination, next hop to the destination and feasible distance are validated. The probability of disjoint is computed for the nodes in order to construct the disjoint routes with minimal loss.

Route Discovery

The sensor nodes that have the aggregated multimedia content then begin to transmit the data through best routing paths, which are also reliable. In order to discover the routes, the nodes broadcast the route request (RREQ) packets to the other nodes. The nodes, which can also perform as routers reply or to RREP packets that contain the sequence number of the nodes. On analysing the received packets, those nodes with the highest sequence numbers are chosen for routing. In order to visualize this concept, the nodes must know the coordinates of the other nodes, which can be detected effectively using the energy efficient node localization algorithm.

Route Selection

It is important that the constructed paths with loop-free and disjoint properties are ranked in terms of better QoS values. The QoS parameters namely energy drain rate, path reliability, average delay and link quality are computed for the constructed paths. The energy drain rate of a node in a route can be calculated by estimating the energy consumed during transmission Et, the energy consumed during reception Er, the energy spent when the nodes are in idle state Ei, and the energy spent for data sensing Es. This energy drain rate Ed is given as follows

$$E_d = E_t + E_r + E_i + E_s \quad (1)$$

Distance D between a selected node and the destination node is computed as follows:

$$D = d[i,j] + length\ [i,j] \quad (2)$$

Where $d[i,j]$ and $length\,[i,j]$ refer to the distance and length between node i and its neighbouring node j.

Path reliability (R_p) is the sum of packet loss rate and the bit error rate (BER) in the path p.

$$R_p = Packet\ loss\ rate + BER \tag{3}$$

Where

$$Packet\ loss\ rate = \frac{Number\ of\ packets\ lost}{Number\ of\ packets\ sent} \tag{4}$$

$$BER = \frac{Number\ of\ errors}{Total\ Number\ of\ bits\ sent} \tag{5}$$

The Average delay (Ad) is given as follows:

$$Average\ delay \quad Ad = \frac{1}{(\mu - \lambda)} \tag{6}$$

Where μ refers to the number of packets handled per second and λ is the average rate at which the packets are arriving at the path.

Link quality () can be estimated in terms of the packets received undamaged in a link during t seconds.

$$Q_L = \frac{Packets\ received\ in\ t}{\max(\ Packets\ expected\ in\ t, Packets\ recieved\ in\ t)} \tag{7}$$

The energy drain rate and distance measures are obtained as in EE-PWDGR and the computed values of the path reliability, average delay and link quality, fuzzy rules are generated and used for selecting the routing paths, based on a priority of the weights calculated with better performance values.

$$Path\ priority = w1 * \frac{1}{Ed} + w2 * \frac{1}{D} + w3 * R_p + w4 * \frac{1}{Ad} + w5 * Q_L \tag{8}$$

Where $\sum w_i = 1$.

Based on the computation results, the paths are ranked in a descending order of routes, with better QoS values ranked higher. The prioritized multimedia contents are now allocated to the paths based on priority. However, in this approach of path ranking, the QoS values of all the routes are compared with one another in order to determine the best path. This comparison process consumes more time, thus increasing the waiting time for data transmission, which in turn could degrade the overall performance of time-constrained

applications. These issues of increased time consumption for path ranking can be minimized by selecting optimal routes using optimization algorithms.

Artificial Bee Colony based Route Optimization

Route selection based on the optimization concept is more efficient than the general comparison based approaches. ABC optimization is a nature-inspired algorithm based on the swarm intelligent honey bee foraging behaviour. The general concept of food search in ABC is the basis for the route optimization in this study. ABC has few important control parameters such as population size, limit, and maximum cycle number. The main advantage of the ABC is its fast convergence speed due to the utilization of the efficient search process in the population initialization phase.

The number of employed bees or onlooker bees is assigned as equal to the number of nodes, with the multimedia data as the input. The routing paths to the destination node are considered as the food sources. The population of solutions is initialized and the employed bees generate a new candidate solution from its neighbourhood such that each employer bee exploits only one solution.

Initialize a random value *rand*. Let SN be the initial population of solutions of routing paths which is equal to the number of bees. The initial population of solution n_i can be found using the equation

$$n_i = n_{min} + rand(0,1) * (n_{max} - n_{min}) \quad (9)$$

Where $i \in (1,2,....SN)$, n_{max} and n_{min} are the lower and upper bounds of n_i

As each employed bee N_i generates new routing paths based on a solution search equation, the solution search equation of ith bee can be given as u_i.

$$u_i = n_i + \emptyset_i * (n_i - n_k) \quad (10)$$

Where $k \in (1,2,....SN)$ is the random index and also such that k is different from i. $\emptyset_i$ is a random number in the range of [-1,1].

Using this equation, the employed bee generates new routing solutions and compares then with the available solutions in memory. If the new routing solution is equal to the QoS values or better QoS values than the old one, the old solution is replaced by the new routing solution. If not, the old solution is retained and the corresponding information is shared with the onlooker bee. The onlooker bee then selects the routing solution based on the QoS values, depending on the probability associated with each solution. The fitness values are calculated based on the weight values of QoS parameters. The probability associated with each routing solution is given as follows:

$$pi = \frac{fitness_i}{\sum_{i=1}^{n} fitness_i} \quad (11)$$

Where $fitness_i$ is the fitness value of the routing solution that can be obtained by equation (12)

$$fitness_i = \begin{cases} \frac{1}{1+f_i} f_i \geq 0 \\ 1 + |f_i| f_i < 0 \end{cases} \quad (12)$$

Where $fitness_i$ is the objective function value which can be computed as

$$fi = w1 * \frac{1}{Ed} + w2 * \frac{1}{D} + w3 * R_p + w4 * \frac{1}{Ad} + w5 * Q_L \quad (13)$$

Where $\sum w_i = 1$.

If the new routing solution has equal or better QoS values than the old solution, then the new one replaces the old solution. After this process, the employed bees which abandon these routing solutions and look for new solutions by becoming scout bees.

Enhanced Artificial Bee Colony based Route Optimization

Since ABC-based Route Optimization has issues in solution search equation, an enhanced ABC with two improved solution search equation, employed bee and onlooker bee separately. The flow of the enhanced ABC for route optimization is shown in Figure 1. The process begins with setting the control parameters for the enhanced ABC while the node information are obtained from the routing tables. Then the number of possible, energy efficient routes with less probability of attacks and packet loss is identified. Then the bees are initialized to compare the routes with the comparison order changed after each iteration. These bees utilize the two new search equations to detect the paths and update them as per their fitness. The best routing path information is saved while the previous best are updated. Thus the routing process is determined. The solution search equations can be determined by the employed bees. After the initialization of population routing solutions, the employed bees generate new solutions based on the newly improved solution search equation (Karaboga & Akay, 2009) as given by the following equation:

$$u_i = n_{best} + \emptyset_i * (n_{best} - n_i) \quad (14)$$

Where n_{best} is the global best solution and $\emptyset_i$ is a random number in the range of [-1,1]. Thus each employed bee generates new routing solutions and shares the same with the onlooker bees. The use of the global best solution improves the search operation with maximum exploitation.

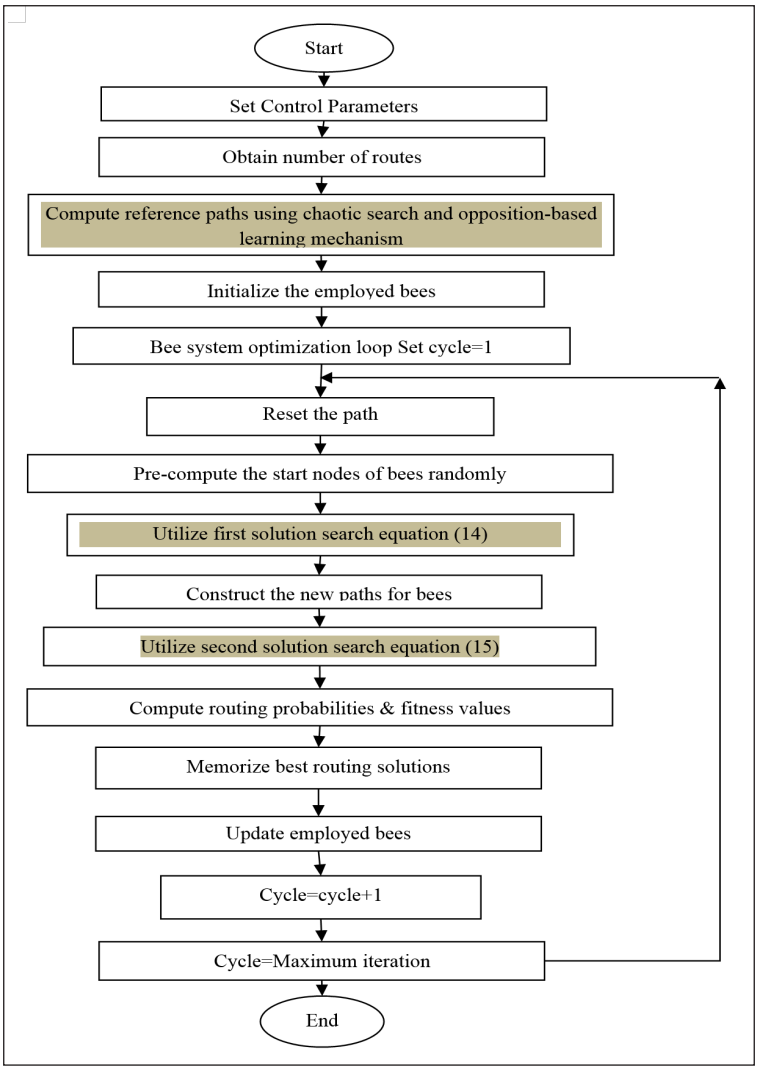


Figure 1. Flow of Enhanced ABC algorithm for Route Optimization

In addition to this equation, yet another improved search equation has also introduced for the search operation in onlookers. The equation makes the best use of the global best solution if the probability values are less than random number *rand*. When the probability using equation (11) is computed, it is compared with the random value initiated at the initialization phase.

$$u_i = n_{best} + \emptyset_i * (n_{best} - n_k) \tag{15}$$

where the index k is randomly selected from (1, 2, . . . , N), and different from the index n_i.

The probabilities of the routing paths are then computed and the best solutions detected. The new candidate solution only around the best solution of the previous iteration is selected using the enhanced ABC algorithm. The optimal routing paths thus obtained are the best routing path. The paths are ranked efficiently based on these paths to allocate the prioritized multimedia content for perfect transmission. This approach further enhances the search operation, results in attaining optimal routes.

In the case of Evolutionary algorithms, random initialization is essentially the most normally used process to generate candidate solutions. However, this affects the convergence speed and the quality of the final solution. To address this issue, the chaotic search algorithm (Alatas, 2010) and opposition-based learning mechanism (Xu, Wang, Wang, Hei, & Zhao, 2014) have been implemented in the initialization phase in order to enhance the solution search.

In chaotic function sinusoidal iterator is chosen and its equation is defined as follows:

$$Ch_{l+1} = \sin(\pi ch_k), \quad Ch_l \in (0,1) \tag{16}$$

Where k is the iteration counter and l is the pre-set maximum number of chaotic iterations.

The random function $\emptyset_i$ used in equation (14) and (15) is now changed as follows for employed and onlooker bees

$$u_i = n_{best} + \text{ch}_{\,l+1} * (n_{best} - n_i) \tag{17}$$

$$u_i = n_{best} + \text{ch}_{\,l+1} * (n_{best} - n_k) \tag{18}$$

Rahnamayan, Tizhoosh and Salama (2008) proposed an opposition based differential evolution (OBDE) to support the convergence speed of DE, which used opposite-based learning (OBL). The proposal of OBL was once offered by Tizhoosh (2005). The same idea can also be utilized in ABC to improve the convergence speed. If the randomized initialized solutions are closer to the global optima, then the required solution can also be determined with much less computational efforts but if the solutions are far from the optima, then it takes higher computational cost or at times, it would also not be infeasible to track the desired solutions. As instructed in these two works, the computational cost could be decreased via making use of the notion of OBL. In OBL, we take into account the easier solutions as good as their opposite counter phase solutions and combine them. The search strategy is then utilized to determine the fittest solutions among them.

$$ou_i = n_{best} + n_i - u_i \tag{19}$$

$$ou_i = n_{best} + n_i - u_i \tag{20}$$

The fittest individuals are found from set u (N) ∪ ou (N) as initial population for both the employed and onlooker bees. If a position update is not improved further for a pre-set number of iteration, then the scout bees discover new solutions ou_i to be replaced.

Procedures for Load Balancing Optimization

In the ABC model of delay-energy trade-off problem, then a number of the employed bees are equal to the number of solutions in the population. Each employed bee is associated with one food source. At the first step, a randomly distributed initial population which, is considered as the initial solution is generated. An onlooker bee chooses a food source depending on the probability value and this probability value P_i is calculated by Equation. (11)

1. Initialize Control Parameters of EABC Algorithm: Colony size CS, dimension of the problem D = 5, and

Limit for scout L = (CS*D)/2.

2. Initialize scout bee positions as per subroutine below:

2.1. Let s be the total no. of scout bees positions.

2.2. Evaluate the fitness of the positioning

3. Set no. of iterations for optimisation, MaxIt.

4. REPEAT for i = 1 to MaxIt

4.1. Assign k employee bee nodes, best out of s scouts, as k_i .

4.2. Produce new solutions for onlooker positions based on Opposition learning, as vij , for each k_i .

4.3 Apply greedy selection process between new nodes.

4.4 $v_{ij} = k_{best}$,i + r_{ij} (k_{best} ,i − k_i)

4.5. Determine the probability values out of their fitness values.

4.6. Produce new solutions for v_{ijout} of this check.

4.7. Record the best solutions yet.

4.8. Check for abandonment condition according to the QoS model,

$$fi = w1 * \frac{1}{Ed} + w2 * \frac{1}{D} + w3 * R_p + w4 * \frac{1}{Ad} + w5 * Q_L$$

& if satisfied replace with new solutions using Opposition based learning.

4.9. Increment i = i + 1.

EABC-EELB-PWDR Algorithm
Initialize nodes N
Get geographical information's of Node i //Node localization
For node i //Route construction

$$P = \sum_{\pi=\pi_i}(1-(1-\alpha)^{|\pi|}) \times \frac{1}{|\pi_i|} \text{ ion} \quad (21)$$

//Where α is the angle $\angle(p,i,j)$ between nodes i and j
Check for minPLT

If ($P(i) \leq 1$&& PLT=minPLT)
Node i$\rightarrow Path\ DR_i$
Else
Node= i+1

End if
End for
For node i //Route Selection
Compute Ed, D,Rp, Ad, Q_L using equation (1) to (7)
Determine path priority based on weight values using Equation (8)

If ($Path\ priority > 1$)
Assign DR_i = high rank
Else If ($0 > Path\ priority \geq 1$)
Assign DR_i =medium rank
Else If ($0 \leq Path\ priority$)
Assign DR_i =low rank

End if
End for
Initialization //EABC-EELB-PWDGR

Generate new candidate routing solutions u_i using Equation (19)

If $(u_i > n_i)$

$n_{ibest} \leftarrow u_i$

Iteration = 1

Else $n_{ibest} \leftarrow n_i$

Iteration = iteration + 1

End if

Compute probability pi using Equation (11)

If $(rand < pi)$

Generate new solutions u_i using Equation (20)

If $(u_i > n_i)$

$n_{ibest} \leftarrow u_i$

Iteration = 1

Else $n_{ibest} \leftarrow n_i$

Iteration = iteration + 1

End if

End for

If iteration =max

Replace n_i with new solution $n_i + 1$ using Equation (9)

End if

$N_i \rightarrow Scout\ bee$

$n_{ibest} \leftarrow DR_i$

Allocate data to DR

End

RESULTS AND DISCUSSION

In this research, the performance of ABC-based EELB-PWDGR and EABC-based EELB-PWDGR have been evaluated using Network Simulator (NS-2) and compared with that of PWDGR, EE-PWDGR, and EELB-PWDGR in terms of end-to-end delay, PSNR, energy per-packet, hop count and network lifetime in order to determine the efficiency of the proposed schemes. The video sensor nodes are utilized for collecting the multimedia data at

different geographical locations with all the nodes considered as static nodes. The utilized sensor node is a 4-layer protocol with satisfactory QoS factors. In the whole network, only a few video monitoring nodes (VN) covering the monitoring area and the capacity of battery equipped for VN node are larger than that of common nodes. The responsibilities of other common nodes are transmitting data to the sink which has infinite energy through many hops. The energy of all the common nodes is equal. The node energy consumption is computed using the sending time, receiving time, idle time and overhearing time for the packet transmission. The NS-2 simulation settings are as followed as utilized in Al-Ariki and Swamy (2017) and Al-Ariki, Mohammed and Swamy (2017a, 2017b) and are given in Table 1.

Figure 2 and Table 2 shows the comparison of PWDGR, EE-PWDGR, EELB-PWDGR, ABC-EELB-PWDGR, and EABC-EELB-PWDGR in terms of End-to-end delay. The delay of EABC-EELB-PWDGR can always be seen to be lower than those by other methods. This becomes the selected path is very close to the bandwidth required by a video stream. Hence, the congestion and data loss are handled based on required QoS in EABC-EELB-PWDGR. It can be seen in Figure 2 and Table 2 that end-to-end delay from 1-Hop to 4-Hop of proposed two approaches is nearly 20% which is lesser than the PWDGR.

Figure 3 and Table 3 shows the comparison in terms of PSNR value. From the figure, it can be observed that ABC-EELB-PWDGR achieves higher PSNR than the EELB-PWDGR, which on an average is about 4dB. EABC-EELB-PWDGR achieves the highest PSNR, which on average is about 8 dB higher than that of EELB-PWDGR for all hop level with varies number of node used for simulation.

Figure 4 and Table 4 presents the comparison in terms of network life of PWDGR, EE-PWDGR, EELB-PWDGR, ABC-EELB-PWDGR and EABC-EELB-PWDGR for 1-Hop, 2-Hop, 3-Hop and 4-Hop level. From the results, it can be concluded that the network lifetime of EABC-EELB-PWDGR and ABC-EELB-PWDGR are significantly longer and stable than EELB-PWDGR in all Hop levels. EABC-EELB-PWDGR determines the best path. Further, the longevity of EABC-EELB-PWDGR is more than ABC-EELB-PWDGR. In analysis of lifetime with varied transmission ranges from 35 to 100, considering for all hop counts the lifetime is increased 17% on an average for EABC-EELB-PWDGR.

Figure 5 and Table 5 shows the comparison in terms of Hop count value. The hop count of ABC-EELB-PWDGR and EABC-EELB-PWDGR is lesser than that of EELB-PWDGR and EE-PWDGR. A hop count is lesser also implies energy consumption is also less. In this case, the average number of hop is reduced 8% for EE-PWDGR, 10% for EELB-PWDGR, 20% for ABC-EELB-PWDGR and 30% for EABC-EELB-PWDGR from PWDGR. EABC-EELB-PWDGR transmits the packets with adaptive distance and optimal path with QoS parameters which in turn reduces the overall hop count.

Figure 6 and Table 6 shows the comparison in terms of Energy per-packet value. Energy per-packet metric is used to measure how much energy is consumed to transmit one packet over a wireless link. The retransmission in the optimally selected path is reduced in the proposed approaches. The energy per-packet in the case of EABC-EELB-PWDGR is almost 60% lesser than PWDGR, while in the case of ABC-EELB-PWDGR approach it is 50% lesser than PWDGR. The scope of this routing protocol can be learnt from the performance improvement. The routing performance is mainly concentrated in terms of delay, power consumption and errors in transmission for which this research utilized five parameters to evaluate the performance. The optimization based routing protocols has the scope to be utilized in any energy constraint environment for various applications consisting of multimedia data.

Table 1
Simulation environment

Network Area (Size)	1000X1000 m^2
Network Topology	Randomized
MAC Layer (IEEE Standard)	IEEE 802.11
Data rate at MAC layer	2 Mbps
Link Layer	LL (Link Layer)
IFQ Type	Queue/DropTail/PriQueue
IFQ Length	50
Antenna Type	Antenna/OmniAntenna
Physical Type	Phy/WirelessPhy
Channel Type	Channel/WirelessChannel
Energy model	**EnergyModel**
Primary energy of common node	0.2 w
Primary energy of video node	5 w
Energy consumption at sending unit time	0.660 w/s
Energy consumption at receiving unit time	0.395 w/s
Energy consumption at overhearing unit time	0.195 w/s
Energy consumption at idle unit time	0.035 w/s
Packet loss rate	0.15%
Simulation Time	350 s
Data Transfer Protocol	TCP/UDP
No. of Sink Nodes	1
No. of Source Nodes	6
No. of Sensor Nodes	30
Maximum Transmission range	100m
Packet Size	2MB
No. of Paths	9

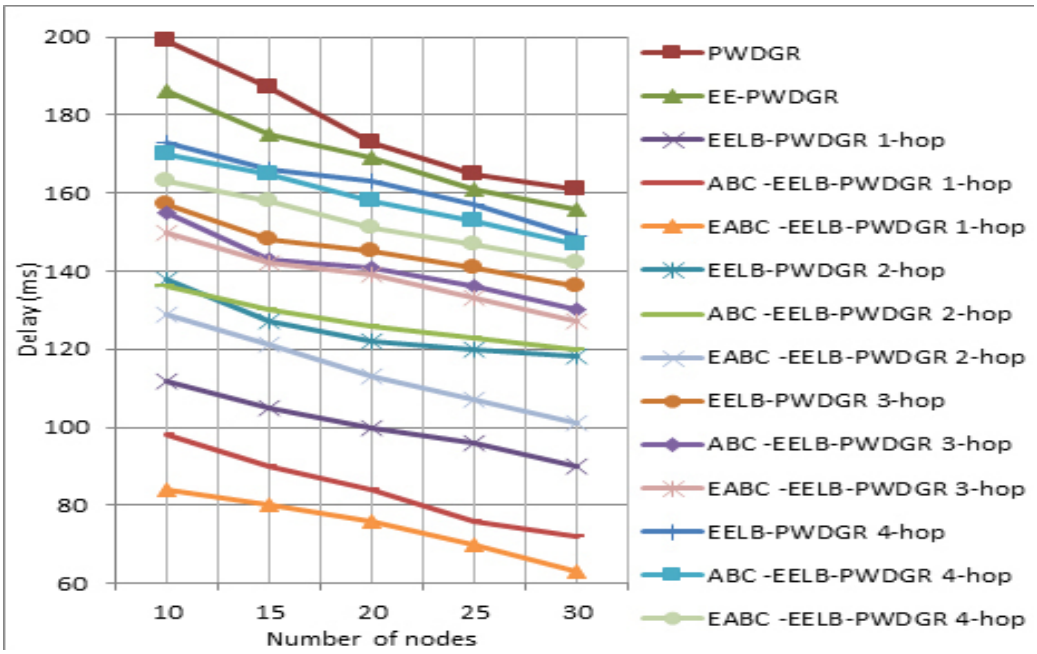


Figure 2. End-to End delay

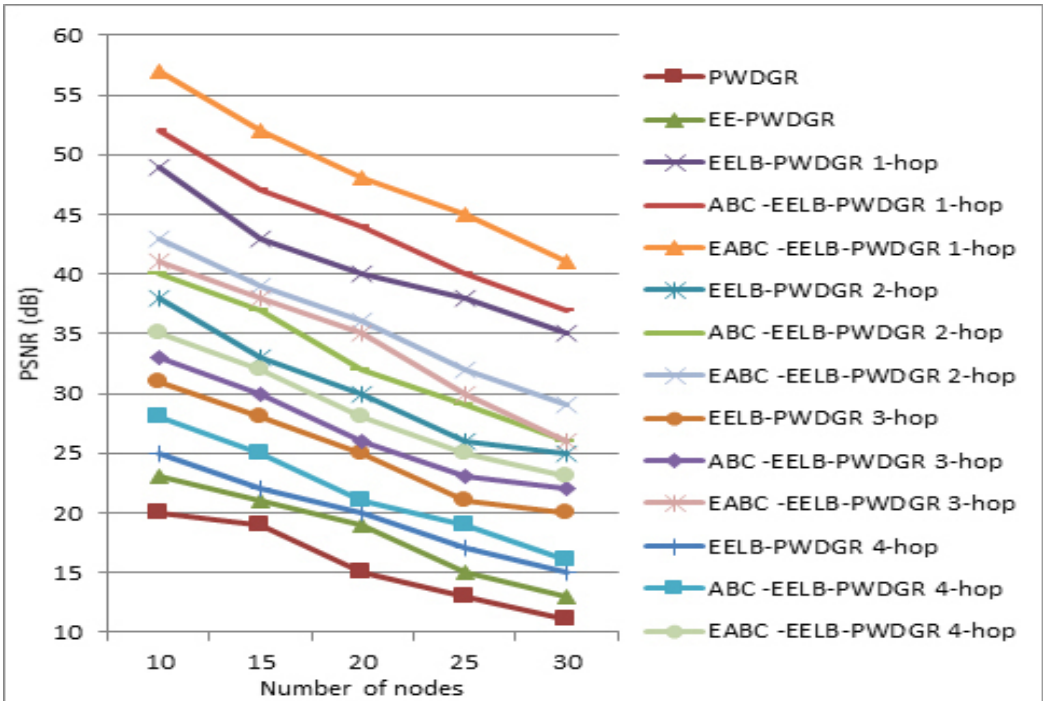


Figure 3. PSNR

Table 2
Comparison in terms of End-to-end delay (ms)

No. of sensor nodes	PWDGR	EE-PWDGR	EELB-PWDGR				ABC -EELB-PWDGR				EABC -EELB-PWDGR			
			1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop
10	199	186	112	138	157	173	98	136	155	170	84	129	150	163
15	187	175	105	127	148	166	90	130	143	165	80	121	142	158
20	173	169	100	122	145	163	84	126	141	158	76	113	139	151
25	165	161	96	120	141	157	76	123	136	153	70	107	133	147
30	161	156	90	118	136	149	72	120	130	147	63	101	127	142

Table 3
Comparison in terms of PSNR (dB)

No. of sensor nodes	PWDGR	EE-PWDGR	EELB-PWDGR				ABC -EELB-PWDGR				EABC -EELB-PWDGR			
			1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop
10	20	23	49	38	31	25	52	40	33	28	57	43	41	35
15	19	21	43	33	28	22	47	37	30	25	52	39	38	32
20	15	19	40	30	25	20	44	32	26	21	48	36	35	28
25	13	15	38	26	21	17	40	29	23	19	45	32	30	25
30	11	13	35	25	20	15	37	26	22	16	41	29	26	23

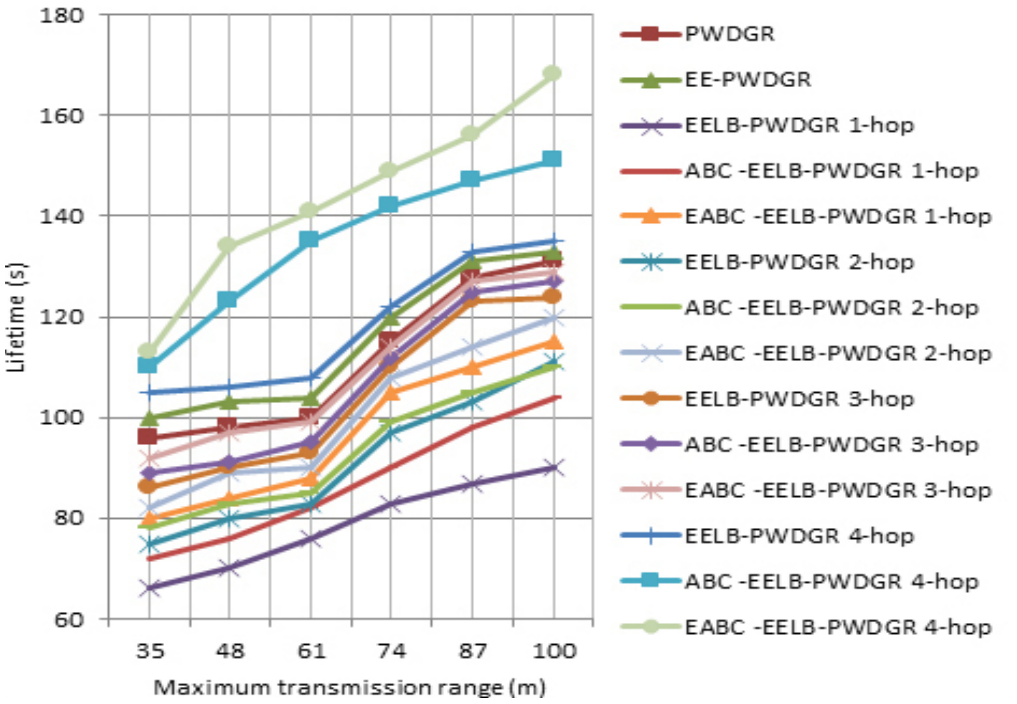


Figure 4. Lifetime

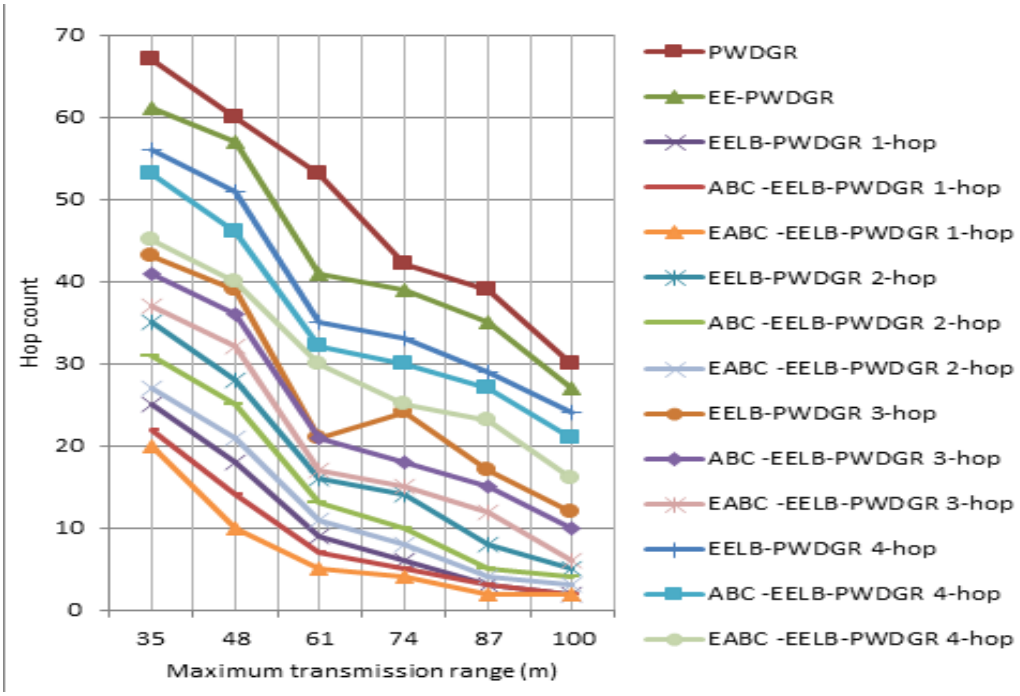


Figure 5. Hop count

Table 4
Comparison in terms of Lifetime

Max transmission range m	PWDGR	EE-PWDGR	EELB-PWDGR				ABC -EELB-PWDGR				EABC -EELB-PWDGR			
			1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop
35	96	100	66	75	86	105	72	78	89	110	80	82	92	113
48	98	103	70	80	90	106	76	83	91	123	84	89	97	134
61	100	104	76	83	93	108	82	85	95	135	88	90	99	141
74	115	120	83	97	110	122	90	99	112	142	105	108	114	149
87	128	131	87	103	123	133	98	105	125	147	110	114	127	156
100	131	133	90	111	124	135	104	110	127	151	115	120	129	168

Table 5
Comparison in terms of Hop count

Max transmission range m	PWDGR	EE-PWDGR	EELB-PWDGR				ABC -EELB-PWDGR				EABC -EELB-PWDGR			
			1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop	1-hop	2-hop	3-hop	4-hop
35	67	61	25	35	43	56	22	31	41	53	20	27	37	45
48	60	57	18	28	39	51	14	25	36	46	10	21	32	40
61	53	41	9	16	21	35	7	13	21	32	5	11	17	30
74	42	39	6	14	24	33	5	10	18	30	4	8	15	25
87	39	35	3	8	17	29	3	5	15	27	2	4	12	23
100	30	27	2	5	12	24	2	4	10	21	2	3	6	16

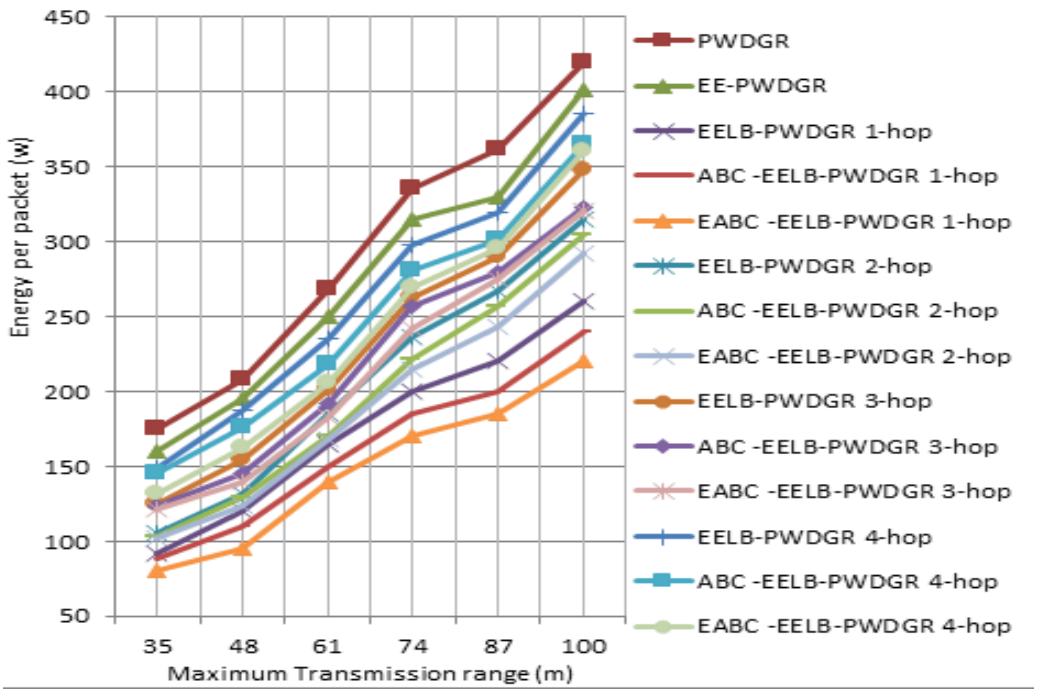


Figure 6. Energy per packet

Table 6
Comparison in terms of Energy per packet (w/10^6)

Max transmission range m	**PWDGR**	**EE-PWDGR**	**EELB-PWDGR**				**ABC -EELB-PWDGR**				**EABC -EELB-PWDGR**			
			1-hop	**2-hop**	**3-hop**	**4-hop**	**1-hop**	**2-hop**	**3-hop**	**4-hop**	**1-hop**	**2-hop**	**3-hop**	**4-hop**
35	175	160	92	105	125	149	88	103	124	145	80	102	121	132
48	208	195	120	132	154	188	110	129	145	176	95	124	140	162
61	268	250	165	185	201	235	150	170	192	218	140	168	183	206
74	335	315	200	236	263	298	185	222	257	281	170	215	242	269
87	361	330	220	267	290	319	200	257	280	301	185	243	275	295
100	419	401	260	315	348	385	240	305	323	365	220	292	320	360

CONCLUSION

Route optimization techniques reduce time consumption to rank routing paths based QoS priority. Hence, this study employs the optimization algorithms ABC and Enhanced ABC with EELB-PWDGR for reducing both latency and energy consumption based on QoS-based load balancing between paths. The Enhanced ABC algorithm consisting of two new improved solutions search equations and two functions for population initialization were used to further reduce the search time and to increase the convergence of ABC. The comparison of the proposed routing schemes shows that the EABC-EELB-PWDGR outperforms the other methods with a 20% less delay, 60% lesser energy consumption and a 17% higher lifetime. The multimedia content quality measured in terms of PSNR is also high for EABC-EELB-PWDGR than that of the approaches. Future research can focus on exploring the performance of EABC-EELB-PWDGR with node fault detection and correction. The event of node failures can significantly affect the overall process of routing and hence automatic fault detection and correction schemes are planned to be implemented along with these routing protocols in future.

ACKNOWLEDGEMENTS

The authors wish to acknowledge J.S.S Research Foundation, J.S.S Technical Institutions Campus, S.J.C.E, Mysore-570006, Karnataka, India for all the facilities provided for this research work.

REFERENCES

Al-Ariki, H. D. E., & Swamy, M. S. (2017). A survey and analysis of multipath routing protocols in wireless multimedia sensor networks. *Wireless Networks*, *23*(6), 1823-1835.

Al-Ariki, H. D. E., Mohammed, A. S., & Swamy, M. S. (2017a). Energy-Efficient Geographical Multi-path Routing Protocol with Adaptive Load Balancing for Wireless Multimedia Sensor Networks. *Journal of Advanced Research in Dynamical and Control Systems, 9*(9), 148-166.

Al-ariki, H. D. E., Mohammed, A. S., & Swamy, M. S. (2017b). A novel approach to energy-efficiency in geographical multi-path routing protocol with adaptive load balancing for wireless multimedia sensor networks. *International Journal of Engineering Intelligent Systems for Electrical Engineering and Communication, 25*(2), 77-92.

Alatas, B. (2010). Chaotic bee colony algorithms for global numerical optimization. *Expert Systems with Applications*, *37*(8), 5682-5687.

Ari, A. A. A., Yenke, B. O., Labraoui, N., Damakoa, I., & Gueroui, A. (2016). A power efficient cluster-based routing algorithm for wireless sensor networks: Honeybees swarm intelligence based approach. *Journal of Network and Computer Applications*, *69*, 77-97.

Azharuddin, M., & Jana, P. K. (2016). Particle swarm optimization for maximizing the lifetime of wireless sensor networks. *Computers & Electrical Engineering*, *51*, 26-42.

Cai, X., Duan, Y., He, Y., Yang, J., & Li, C. (2015). Bee-sensor-C: an energy-efficient and scalable multipath routing protocol for wireless sensor networks. *International Journal of Distributed Sensor Networks*, *11*(3), 0-14.

Chatterjee, S., & Das, S. (2015). Ant colony optimization based enhanced dynamic source routing algorithm for mobile Ad-hoc network. *Information Sciences*, *295*, 67-90.

Ding, W., Tang, L., & Ji, S. (2016). Optimizing routing based on congestion control for wireless sensor networks. *Wireless Networks*, *22*(3), 915-925.

Gupta, G. P., & Jha, S. (2018). Integrated clustering and routing protocol for wireless sensor networks using Cuckoo and Harmony Search based metaheuristic techniques. *Engineering Applications of Artificial Intelligence*, *68*, 101-109.

Hu, X. M., & Zhang, J. (2010). Ant routing optimization algorithm for extending the lifetime of wireless sensor networks. In *Proceedings of 2010 IEEE International Conference on Systems Man and Cybernetics (SMC)* (pp. 738-744). Istanbul, Turkey.

Karaboga, D., & Akay, B. (2009). A comparative study of artificial bee colony algorithm. *Applied mathematics and Computation*, *214*(1), 108-132.

Le Dinh, T., & Nguyen, D. T. (2010). Greedy Geographic Routing with Path Optimization in Wireless Sensor Networks. In *Proceedings of IEEE RIVF International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)* (pp. 1-6). Hanoi, Vietnam.

Paschalidis, I. C., & Li, B. (2011). Energy optimized topologies for distributed averaging in wireless sensor networks. *IEEE Transactions on Automatic Control*, *56*(10), 2290-2304.

Rahnamayan, S., Tizhoosh, H. R., & Salama, M. M. (2008). Opposition-based differential evolution. *IEEE Transactions on Evolutionary Computation*, *12*(1), 64-79.

Rao, P. S., & Banka, H. (2017). Novel chemical reaction optimization based unequal clustering and routing algorithms for wireless sensor networks. *Wireless Networks*, *23*(3), 759-778.

Selvi, B. K., & Latha, S. (2015). Conservatory of lifespan of heterogeneous wireless sensor network using ABC algorithm. *Indian Journal of Science and Technology*, *8*(31), 1-5.

Sirdeshpande, N., & Udupi, V. (2017). Fractional lion optimization for cluster head-based routing protocol in wireless sensor network. *Journal of the Franklin Institute*, *354*(11), 4457-4480.

Sohan, R., Mittal, N., Singh, U., & Sohi, B. S. (2018). An Optimal Tree-Based Routing Protocol Using Particle Swarm Optimization. In *Nature Inspired Computing* (pp. 117-124). Singapore: Springer.

Srie Vidhya Janani, E., & Ganesh Kumar, P. (2015). Energy efficient cluster based scheduling scheme for wireless sensor networks. *The Scientific World Journal*, *2015*, 1-9.

Tizhoosh, H. R. (2005). Opposition-based learning: a new scheme for machine intelligence. In *Proceedings of 2005 International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on Computational Intelligence For Modelling, Control and Automation* (pp. 695-701). Vienna, Austria.

Wang, J., Zhang, Y., Wang, J., Ma, Y., & Chen, M. (2015). PWDGR: pair-wise directional geographical routing based on wireless sensor network. *IEEE Internet of Things Journal*, *2*(1), 14-22.

Wang, L., & Xu, Q. (2010). GPS-free localization algorithm for wireless sensor networks. *Sensors*, *10*(6), 5899-5926.

Xiang, W., Wang, N., & Zhou, Y. (2016). An energy-efficient routing algorithm for software-defined wireless sensor networks. *IEEE Sensors Journal*, *16*(20), 7393-7400.

Xu, Q., Wang, L., Wang, N., Hei, X., & Zhao, L. (2014). A review of opposition-based learning from 2005 to 2012. *Engineering Applications of Artificial Intelligence*, *29*, 1-12.

Ye, Z., & Mohamadian, H. (2014). Adaptive clustering based dynamic routing of wireless sensor networks via generalized ant colony optimization. *IERI Procedia*, *10*, 2-10.

Zeng, B., & Dong, Y. (2016). An improved harmony search based energy-efficient routing algorithm for wireless sensor networks. *Applied Soft Computing*, *41*, 135-147.

Zhang, S., & Liu, S. (2015). Artificial bee colony algorithm with improved search equations. *Journal of Information & Computational Science*, *12*(10), 4069-4076.

Zheng, W., & Luo, D. (2014). Routing in wireless sensor network using Artificial Bee Colony algorithm. In *Proceedings of 2014 International Conference on Wireless Communication and Sensor Network (WCSN)* (pp. 280-284). Wuhan, China.

Pertanika J. Sci. & Technol. 26 (4): 1975 - 1998 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Performance Characteristics and Thermodynamic Investigations on Single-Stage Thermoelectric Generator and Heat Pump Systems

Rajesh Arora[1]* and Ranjana Arora[2]

[1]*Department of Mechanical Engineering, Amity University Haryana, Gurgaon-122413, India*
[2]*Renewable Energy Department, Amity University Haryana, Gurgaon-122413, India*

ABSTRACT

The thermodynamic analysis of thermoelectric devices (TEs) discards the impact caused by heat leak between source and sink. It could lead towards the partial/incomplete modelling of TEs along with some analytical gaps in their performance evaluation. Conversely, appropriate agreement among different design constraints of TEs is a must to upgrade their operating characteristics. In view of this, the modelling of multi-element single-stage thermoelectric generator (TEG) and thermoelectric heat pump (TEHP) is carried out in matrix laboratory (MATLAB 9.2). The irreversibilities caused by heat leak between the source/sink along with Fourier/Joule effects are considered for the modelling and analysis. The power output/thermal efficiency and heating capacity/coefficient of performance (COP) of TEs are analytically derived and optimized on the basis of finite time thermodynamic principles. The predetermined thermoelectric couples are chosen to maximize the heating capacity/COP in proposed configurations. Moreover, the influence of design variables viz electrical current, thermoelectric couples on system throughput is analyzed and presented. The effects of geometrical parameters viz length and area of individual modules on the performance of TEG and TEHP are also discussed.

Keywords: Finite time thermodynamic, thermoelectric generators, thermoelectric heat pumps, (FTT), thermodynamic optimization

ARTICLE INFO

Article history:
Received: 07 February 2018
Accepted: 28 May 2018
Published: 24 October 2018

E-mail addresses:
rajesharora1219@rediffmail.com (Rajesh Arora)
ranjana1219@rediffmail.com (Ranjana Arora)
* Corresponding author

INTRODUCTION

In today's era, the rapid rise in energy demand has led researchers/scientists to search for new resources of energy. As the conventional energy resources are coming to an end, innovative/smart energy conversion technologies are gaining wide

ISSN: 0128-7680
e-ISSN: 2231-8526

attraction globally. Due to non-moving parts, low maintenance requirements, low/zero noise operation, compact configuration and small volume occupancy, these are getting very popular amongst various commercial applications. The thermoelectric devices viz. thermoelectric generators and heat pumps are direct thermal to electricity conversion devices requiring low maintenance, capable to operate in modular form and have possibility to utilize heat from sources over a wide range of temperature (Angrist, 1994; Disalvo, 1999; Goldsmith, 1964; Rowe, 2005). These devices consist of a semiconductor thermocouple having two junctions. When these two junctions, named as hot junction and cold junction are maintained at different temperatures, the voltage is generated in the order of micro-volts/Kelvin. This is known as Seebeck effect. The reverse effect in which a junction can be heated or cooled with the passage of current through it, is known as Peltier effect. Several thermocouples are joined in electrically series configuration to have desired voltage output (Chen & Wu, 2000; Min et al., 2003; Riffat & Ma, 2003; Xuan et al., 2002). Thermoelectric devices are coupled with the source/sink and according to classical thermodynamics, it takes infinite time for heat transfer between system and coupled thermal reservoirs. This is not desirable from practical point of view. The finite time heat transfers are the need of practical thermoelectric devices to obtain finite output power/thermal efficiency, heating capacity/ coefficient of performance (Andresen, 1983). It is observed that classical thermodynamic does not provide the comprehensive picture of equilibrium states/reversible processes. On the other hand, finite time thermodynamic (FTT) provides an extension to conventional thermodynamic theory as it accounts for finite time consumption of heat transfer between system and coupled thermal reservoirs. Finite time thermodynamic approach takes into account both external and internal irreversibilities of the system. In thermoelectric devices, internal irreversibility is because of Joule/conduction heat loss while external irreversibility is because of finite rate heat transfers between heat reservoirs/system (Chen et al., 1994). In the present literature, several investigations are carried out on thermodynamic optimization and analysis of thermoelectric generators/heat pumps with conventional/non-equilibrium thermodynamics. Chen et al. (2005a) enunciated optimization of series mode operated multi-stage TEG with respect to different designing operators. Afterwards, they investigated the characteristics of thermoelectric refrigerator (Chen et al., 2005b). Later, Chen et al. (2008) optimized the throughput of 2-stage heat pumps in electrically series mode with internal/external irreversibilities employing finite time thermodynamic (FTT) principles. David et al. (2012) investigated thermodynamic studies and optimizations of TEHP in view of different design/operating variables and designed a heat exchanger for a system. Afterwards, an experimental model of TEGs had been tested in order to characterize the different performance parameters (Chen et al., 2012; Zhu et al., 2011). A comprehensive comparison of single/two-stage series connected TEGs operating with an automobile exhaust was demonstrated by Sun et al. (2014). They observed that two-stage system can

yield better output in case of high source-side temperature. Riffat et al. (2004) developed the strategy to optimize thermoelectric modules for space conditioning applications where as the optimum design of the modules for space heating applications in the buildings was presented by Kim et al. (2014). Kaushik et al. (2015) analyzed single-stage thermoelectric heat pumps (TEHP) in view of energy/exergy concepts. They noticed that the effects of internal irreversibility were more in comparison with external irreversibility. Hans et al. (2015) enunciated an optimization study for series/ parallel modes operated TEGs, with respect to various design variables for obtaining maximum output power and first law efficiency. Later, they carried out the comparative performance optimization of 2-stage thermoelectric heat pumps in three different modes (Hans et al., 2016). Further, Kumar and Pandey (2017) calculated the efficiency of a TEG in order to study the impacts of different thermoelectric materials.

In the present literature, a few thermodynamic investigations and optimization studies are available on multi-element single-stage thermoelectric generators and heat pumps (Hans et al. 2015; Nami et al. 2017; Tan et al. 2017). Moreover, the influence of heat leak between the source and sink has not been considered which leads to incomplete modelling of the system. Additionally, there is a need of analyzing the effects of geometric variables and temperature dependent physical properties on performance parameters and optimal designing of thermoelectric devices. In this context, the thermodynamic models of multi-element irreversible TEG and TEHP systems taking into account both internal and external effects, with and without heat leak between source and sink are developed in MATLAB. The performance analysis of above mentioned devices is done on the basis of Newton's law of heat transfer and finite time thermodynamic principles. A comparative study of multi-element irreversible TEG and TEHP based on finite time thermodynamic modeling is carried out considering internal/ external effects. A formulation of performance parameters of thermoelectric devices with respect to electric current, thermoelectric couples, temperature of hot/cold side and thermal conductances of hot/cold side is presented. The internal irreversibilities are caused by Joule/conduction heat losses while, external irreversibilities are because of finite heat transfers between reservoir and devices. The impacts of design factors on output power/thermal efficiency and heating capacity/ coefficient of performance (COP) are analyzed and outcomes are presented graphically. The effects of geometric parameters viz. length and area of thermoelectric elements have been investigated on performance aspects of above mentioned systems. The geometry of the thermoelectric element is directly linked to thermoelectric material and cost of manufacturing. The major outcomes of this study are the comparative investigations along with novel analytical and parametric analysis of multi-element thermoelectric systems with and without heat leak considerations.

Finite-Time Thermodynamic Modeling of Thermoelectric Devices

Thermodynamic modeling of multi-element single-stage thermoelectric generator (TEG) and thermoelectric heat pumps (TEHP) with both external/internal irreversibilities is done through finite-time thermodynamic principles. The thermodynamic modeling of the proposed TEs has been carried out based on the following assumptions.

- Steady-state one-dimensional heat flow has been considered.
- The exoreversible configurations with negligible contact/interconnection resistances are considered.
- The Thomson effect in thermoelectric elements is not taken into account.
- Convection and radiation losses in thermoelectric elements are not considered.
- The material used for TECs is assumed to be Bismuth Telluride (Bi_2Te_3).

The following temperature reliant properties of the thermoelectric materials, as given by Xuan et al. (2002) are used in the present study.

$$\alpha = \left[\alpha_p - (-\alpha_n)\right] = 2 \times (22224.0 + 930.6T_{av} - 0.9905T_{av}^2) \times 10^{-9} V / K \quad (1)$$

$$\rho_n = \rho_p = (5112.0 + 163.4T_{av} + 0.6279T_{av}^2) \times 10^{-10} \Omega / m \quad (2)$$

$$k_n = k_p = (62605.0 - 277.7T_{av} + 0.413T_{av}^2) \times 10^{-4} W / mK \quad (3)$$

$$\tau = \left[\tau_p - (-\tau_n)\right] = 2 \times (930.6T_{av} - 1.981T_{av}^2) \times 10^{-9} V / K \quad (4)$$

$$T_{av} = \left(\frac{T_h + T_c}{2}\right) \quad (5)$$

$$R = \left(\frac{\rho_p L_p}{A_p} + \frac{\rho_n L_n}{A_n}\right) \quad (6)$$

$$K = \left(\frac{k_p A_p}{L_p} + \frac{k_n A_n}{L_n}\right) \quad (7)$$

Thermodynamic Modeling of Single-Stage Thermoelectric Generator (TEG)

The TEG model with P/N type semiconductor legs is illustrated in Figure 1(a). The thermoelectric couples are arranged in electrical series and thermally parallel configurations in pursuance of generating a considerable voltage output from the system. The generalized TEG model revealing the heat flow and generating the electrical output is presented in Figure 1(b). In corresponding figures, T_H and T_h signifies the input source/hot junction

temperatures whereas the T_L and T_c are the heat sink and cold junction ones. Larger the differential temperature amongst the two junctions, more is the potential difference generated at external resistance R_L.

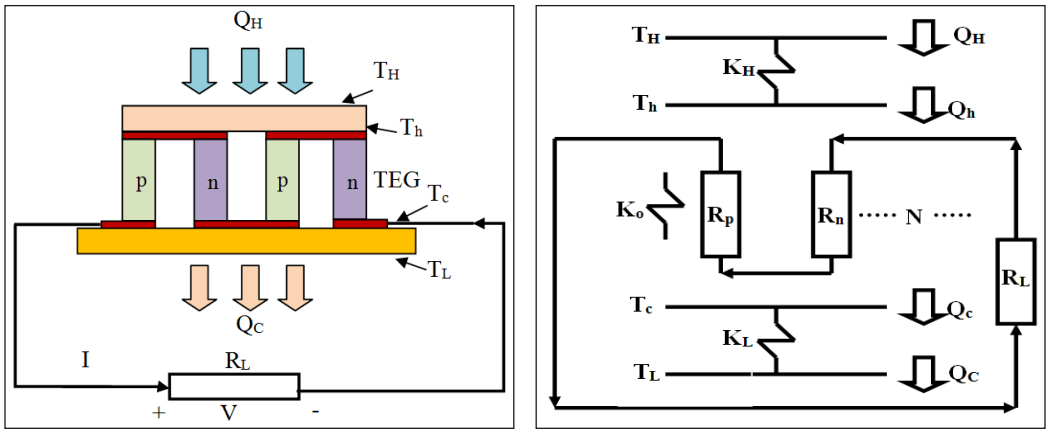


Figure 1. (a) Single-stage TEG; and (b) Generalized FTT model of TEG

The energy equations for P-type semiconductor legs can be written as

$$Q_{Kin}\text{-}Q_{Kout}+Q_J=0 \qquad (8)$$

where, Q_{Kin} and Q_{Kout} are the heat input/output of the thermoelectric elements by conduction and Q_J is the Joule heat generated in the element. Equation 8 can be rewritten as

$$\frac{d}{dx}\left(T_P+dT_P\right)k_P A_P-\frac{dT_P}{dx}k_P A_P+(J_P)^2\frac{A_P L_P dx}{\sigma_P L_P}=0 \qquad (9)$$

Similarly, the equation for N-type semiconductor leg can be stated as

$$\frac{d}{dx}\left(T_N+dT_N\right)k_N A_N-\frac{dT_N}{dx}k_N A_N+(J_N)^2\frac{A_N L_N dx}{\sigma_N L_N}=0 \qquad (10)$$

Boundary conditions for P/N types semiconductor leg can be given as

$$T_P(0)=T_N(0)=T_c \qquad (11)$$

$$T_P(L_P)=T_N(L_P)=T_h \qquad (12)$$

The temperature dependent thermoelectric properties k_P, σ_P are the functions of T_P and k_N, σ_N are the functions of T_N. Putting the boundary conditions to Equations (9) and (10) and substituting k, σ by their mean values $\overline{k}$ and $\overline{\sigma}$, differential equations can be given as

$$\overline{k}_P A_P\frac{d^2T_P}{dx^2}+\frac{I^2}{\overline{\sigma}_P A_P}=0 \qquad (13)$$

$$\overline{k}_N A_N \frac{d^2 T_N}{dx^2} + \frac{I^2}{\overline{\sigma}_N A_N} = 0 \tag{14}$$

The total heat flows via the hot-side/cold-side junctions are

$$Q_h = N\left(\left(\alpha_{Ph} - \alpha_{Nh}\right) T_h I + k_P\big|_{x=L_P} A_P \frac{dT_P}{dx}\bigg|_{x=L_P} + k_N\big|_{x=L_N} A_N \frac{dT_N}{dx}\bigg|_{x=L_N} \right) \tag{15}$$

$$Q_c = N\left(\left(\alpha_{Pc} - \alpha_{Nc}\right) T_c I + k_P\big|_{x=0} A_P \frac{dT_P}{dx}\bigg|_{x=0} + k_N\big|_{x=0} A_N \frac{dT_N}{dx}\bigg|_{x=0} \right) \tag{16}$$

The rate of heat flows from heat source/sink are given as

$$Q_H = K_H (T_H - T_h) \tag{17}$$

$$Q_C = K_L (T_c - T_L) \tag{18}$$

The heat flows from hot/cold junctions are given as

$$Q_h = N[\alpha_h I T_h + k(T_h - T_c) - 0.5 I^2 R] \tag{19}$$

$$Q_c = N[\alpha_c I T_c + k(T_h - T_c) + 0.5 I^2 R] \tag{20}$$

Where,

$\alpha_h = \alpha_{hp} - \alpha_{hn}$ and $\alpha_c = \alpha_{ch} - \alpha_{cp}$

According to heat flow balance, one can have following condition as

$$Q_H = Q_h \tag{21}$$

$$Q_C = Q_c \tag{22}$$

The output is given by

$$P = Q_h - Q_c \tag{23}$$

By putting the values of Q_h and Q_c in Eq. (23) from Equations (19) and (20)

$$P = NI(\alpha_h T_h - \alpha_c T_c - IR) \tag{24}$$

The efficiency is given by

$$\eta = \frac{P}{Q_h} = \frac{NI[\alpha_h T_h - \alpha_c T_c - IR]}{N[\alpha_h I T_h + k(T_h - T_c) - 0.5 I^2 R]} \tag{25}$$

Putting, $\alpha_h = \alpha_c = \alpha$

From above Equations, one can get T_h and T_c as

$$T_h = \frac{0.5N^2\alpha RI^3 + (-N^2Rk - 0.5NRK_L)I^2 + N\alpha K_H T_H I - NK(K_H T_H + K_L T_L) - K_L K_H T_H}{N^2\alpha^2 I^2 + N\alpha I(K_H - K_L) - NK(K_H + K_L) - K_L K_H} \quad (26)$$

$$T_c = \frac{0.5N^2\alpha RI^3 + (N^2Rk + 0.5NRK_H)I^2 + N\alpha K_L T_L I + NK(K_H T_H + K_L T_L) + K_L K_H T_L}{-N^2\alpha^2 I^2 + N\alpha I(K_L - K_H) + NK(K_H + K_L) + K_L K_H} \quad (27)$$

TEG Model with Heat Leak between Source and Sink

The heat flows from hot/cold junctions due to small heat leak is given as

$$Q_c^{'} = N[\alpha_c I T_c + k(T_h - T_c) + 0.5I^2R + K_0(T_c - T_L)] \quad (28)$$

$$Q_c^{'} = N[\alpha_c I T_c + k(T_h - T_c) + 0.5I^2R + K_0(T_c - T_L)] \quad (29)$$

The output is given as

$$P = Q_h^{'} - Q_c^{'} \quad (30)$$

The efficiency in this case can be given as

$$\eta = \frac{P}{Q_h^{'}} = \frac{NI[\alpha_h T_h - \alpha_c T_c - IR]}{N[\alpha_h I T_h + k(T_h - T_c) - 0.5I^2R + K_0(T_H - T_h)]} \quad (31)$$

The values of various parameters used in FTT modeling of multi-element TEG is given in Table 1.

Table 1
Value (s) of various parameters

Parameters	Values
R	2.885×10^{-3} Ω
K	0.02772 W/K
α	425×10^{-6} V/K
T_H	400 K
T_L	300 K
N	127
K_H=K_L	10W/K

Source: (Chen et al., 2005)

Thermodynamic Modeling of Single-Stage Thermoelectric Heat Pump (TEHP)

The TEHP model with P/N type semiconductor legs is illustrated in Figure 2(a). The thermoelectric couples are arranged in electrical series and thermally parallel configurations. The generalized TEHP model revealing the heat flow and with the electrical input is presented in Figure 2(b). In corresponding figures, T_H and T_h signifies the input source/ hot junction temperatures whereas the T_L and T_c are the heat sink and cold junction ones.

When electrical current, 'I' flows through TEHP, heat flows $N\alpha IT_c$ and $N\alpha IT_h$ are captivated from the surroundings and rejected to the space to be heated. The heat source is maintained at low temperature whereas the sink i.e., heated space is maintained at high temperature.

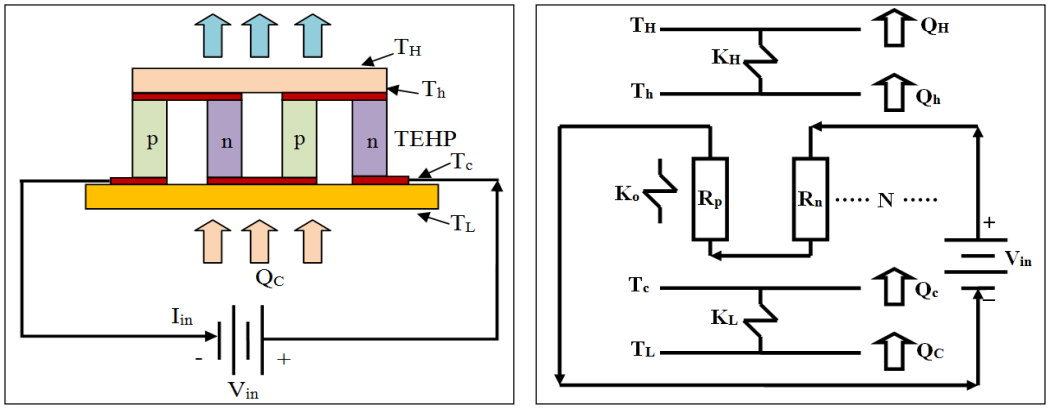


Figure 2. (a) Single stage TEHP; and (b) Generalized FTT model of TEHP

The energy equations of P-type semiconductors are given as

$$Q_{Kin}-Q_{Kout}+Q_J=0 \tag{32}$$

where, Q_{Kin} and Q_{Kout} are the heat input/output of the thermoelectric element by conduction and Q_J is the Joule heat generated in the element. Eq. 32 can be rewritten as

$$\frac{d}{dx}\left(T_P+dT_P\right)k_PA_P-\frac{dT_P}{dx}k_PA_P+(J_P)^2\frac{A_PL_Pdx}{\sigma_PL_P}=0 \tag{33}$$

Similarly, the equation for N-type semiconductor leg can be written as

$$\frac{d}{dx}\left(T_N+dT_N\right)k_NA_N-\frac{dT_N}{dx}k_NA_N+(J_N)^2\frac{A_NL_Ndx}{\sigma_NL_N}=0 \tag{34}$$

Boundary conditions for P and N type semiconductor legs can be given as

$$T_P(0)=T_N(0)=T_c \tag{35}$$

$$T_P(L_P) = T_N(L_P) = T_h \tag{36}$$

The temperature dependent thermoelectric properties k_P, σ_P are the functions of T_P and k_N, σ_N are the functions of T_N. Putting the boundary conditions to Equations (33) and (34) and substituting k, σ by their mean values $\overline{k}$ and $\overline{\sigma}$, differential equations can be given as

$$\overline{k}_N A_N \frac{d^2T_N}{dx^2} + \frac{I^2}{\overline{\sigma}_N A_N} = 0 \tag{37}$$

$$\overline{k}_N A_N \frac{d^2T_N}{dx^2} + \frac{I^2}{\overline{\sigma}_N A_N} = 0 \tag{38}$$

The total heat flows via the hot-side/cold-side junctions are

$$Q_h = N\left(\left(\alpha_{Ph} - \alpha_{Nh}\right) T_h I - k_P\big|_{x=L_P} A_P \left.\frac{dT_P}{dx}\right|_{x=L_P} - k_N\big|_{x=L_N} A_N \left.\frac{dT_N}{dx}\right|_{x=L_N} \right) \tag{39}$$

$$Q_c = N\left(\left(\alpha_{Pc} - \alpha_{Nc}\right) T_c I - k_P\big|_{x=0} A_P \left.\frac{dT_P}{dx}\right|_{x=0} - k_N\big|_{x=0} A_N \left.\frac{dT_N}{dx}\right|_{x=0} \right) \tag{40}$$

The rate of heat flows from heat source/sink are given as

$$Q_H = K_H(T_h - T_H) \tag{41}$$

$$Q_C = K_L(T_L - T_c) \tag{42}$$

The rate of heat flows from hot/cold junctions are given as

$$Q_h = N[\alpha_h I T_h - k(T_h - T_c) + 0.5I^2R] \tag{43}$$

$$Q_c = N[\alpha_c I T_c - k(T_h - T_c) - 0.5I^2R] \tag{44}$$

Putting,

$$\alpha_h = \alpha_c = \alpha$$

According to heat flow balance, one can have following condition as

$$Q_H = Q_h \tag{45}$$

$$Q_C = Q_c \tag{46}$$

The heating power is given by

$$P = Q_h - Q_c \tag{47}$$

The COP of TEHP can be calculated as

$$COP = \frac{Q_h}{Q_h - Q_c} \tag{48}$$

From Eqs. (45) and (46), one can get T_h and T_c as

$$T_h = \frac{0.5N^2\alpha RI^3 + (N^2Rk + 0.5NRK_L)I^2 + N\alpha K_H T_H I + NK(K_H T_H + K_L T_L) + K_L K_H T_H}{-N^2\alpha^2 I^2 + N\alpha I(K_H - K_L) + NK(K_H + K_L) + K_L K_H} \tag{49}$$

$$T_c = \frac{0.5N^2\alpha RI^3 + (-N^2Rk - 0.5NRK_L)I^2 + N\alpha K_H T_H I - NK(K_H T_H + K_L T_L) - K_L K_H T_H}{N^2\alpha^2 I^2 + N\alpha I(K_L - K_H) - NK(K_H + K_L) - K_L K_H} \tag{50}$$

TEHP Model with Heat Leak between Source and Sink

The heat flows from hot/cold junctions due to small heat leak is given as

$$Q_h^{'} = N[\alpha_h I T_h + k(T_h - T_c) - 0.5I^2R - K_0(T_H - T_h)] \tag{51}$$

$$Q_c^{'} = N[\alpha_c I T_c + k(T_h - T_c) + 0.5I^2R - K_0(T_c - T_L)] \tag{52}$$

The heating power is given as

$$P = Q_h^{'} - Q_c^{'} \tag{53}$$

The COP of TEHP with heat leak

$$COP = \frac{Q_h^{'}}{Q_h^{'} - Q_c^{'}} \tag{54}$$

The values of various parameter used in FTT modeling of multi-element TEHP system is given in Table 2.

Table 2
Value (s) of various parameters

Parameters	Values
R	2.885×10^{-3} Ω
K	0.02772 W/K
α	425×10^{-6} V/K
T_H	300 K
T_L	250 K
N	127
K_H=K_L	100 W/K

Source: (Chen et al., 2008)

DISCUSSION

Figure 3 shows the variations of output power (P) with electric current (I) for various values of sink temperature (T_L), assuming a fixed source temperatures (T_H) of multi-element TEG. The output power is a parabolic function of "I" and its peak value decreases as T_L increases. For a fixed value of source temperature, as T_L increases the temperature differences (DTG) between the heat source/sink decreases and power output eventually falls down.

Likewise, the variation of output power (P) with electric current (I) for various values of source temperatures (T_H), assuming a fixed sink temperature (T_L) of TEG is demonstrated in Figure 3. The output power is a parabolic function of I and its maximum value increases with increase in the value of T_H. For a fixed value of sink temperature, as T_H increases the temperature differences (DTG) between heat source and sink increases and power output ultimately goes up.

Figure 5 and Figure 6 shows the parabolic variations of first law efficiency (η) with I for different values of T_H and T_L respectively. The thermal efficiency increases as T_H increases (for fixed value of T_L) and drops with increase in T_L (for fixed value of T_H). The reason behind this is, as the value of DTG increases in TEG system the first law efficiency goes up and vice versa.

The variation of P with I for different number of thermoelectric elements (N) is illustrated in Figure 7. The output power attains its peak at an optimum value of N, nearly 120. Number of thermoelectric elements affects power output of TEG system for optimal values.

The variations in P and η with I for different values of thermal conductance at hot and cold junction (K_H and K_L) are illustrated in Figure 8 and Figure 9. As the value of K_H and K_L increases, maximum values attained by P and η also go up. The power output and thermal efficiency first increases with current but afterwards it decreases due to dominating effect

of Joule internal irreversibility (I^2R). Therefore, one has to obtain the optimum current with respect to desirable power output and thermal efficiency of the system for given set of design variables.

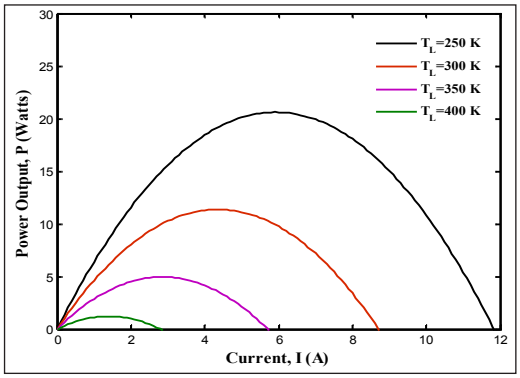


Figure 3. Power output vs 'I' with different values of 'T_L' and fixed value of T_H for TEG

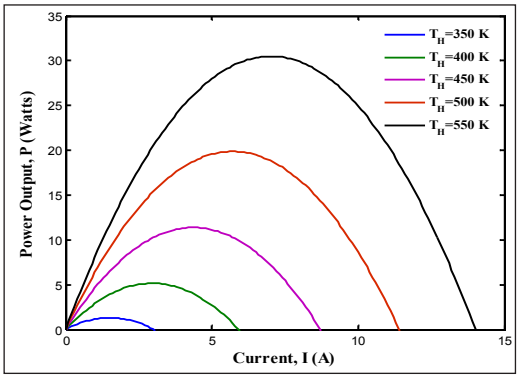


Figure 4. Power output vs 'I' with various values of 'T_H' and fixed value of T_L for TEG

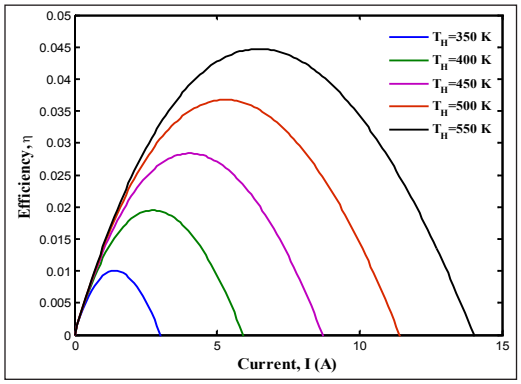


Figure 5. Thermal efficiency vs 'I' with various values of 'T_H' and fixed value of T_L for TEG

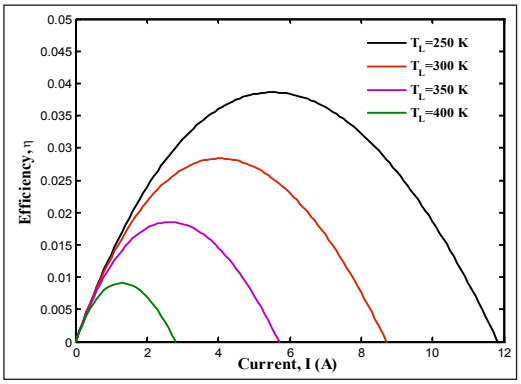


Figure 6. Thermal efficiency vs 'I' with different values of 'T_L' and fixed value of T_H for TEG

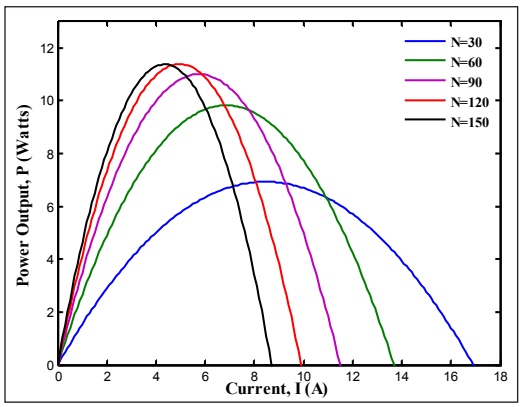


Figure 7. Power output vs 'I' with different values of number of thermoelectric elements for TEG

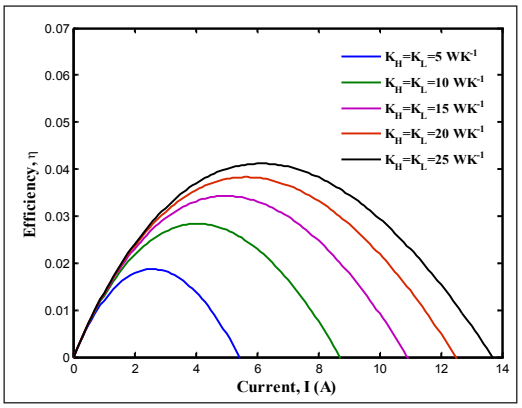


Figure 8. Thermal Efficiency vs 'I' with different values of K_H and K_L for TEG

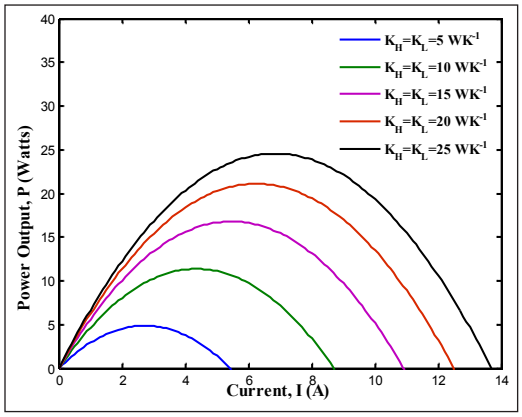


Figure 9. Power Output vs 'I' with different values of K_H and K_L for TEG

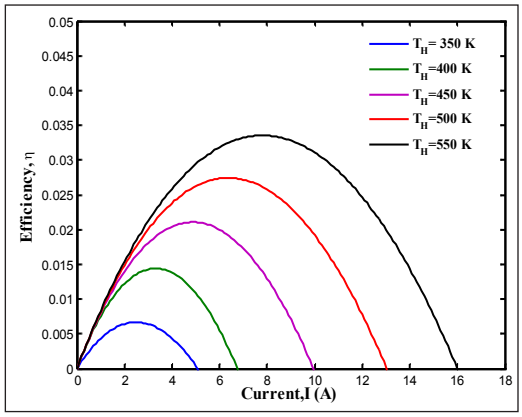


Figure 10. Thermal efficiency vs 'I' with different values of 'T_H' and fixed 'T_L' for TEG with heat leak between source and sink

Figure 10 to Figure 12 show the relationship of thermal efficiency with current for different values of T_H, T_L and thermal conductance (K_H and K_L) with addition of an external irreversibility K_0 in the system. This external irreversibility K_0 is because of heat leak between source/sink, due to which there is considerable drop in the value of maximum thermal efficiency attained by the device.

Figure 13 shows the variations of COP with electric current (I) for various values of sink temperature (T_L), assuming a fixed source temperatures (T_H) of multi-element TEHP. The COP initially increases up to a maximum limit and declines afterwards with working electrical current for all chosen values of T_L. The TEHP yields better COP as the value of sink temperature goes up because for fixed heat source temperature, as the temperature difference between source and sink (DTHP) decreases COP rises ultimately. Similarly, the variation of COP with electric current (I) for different values of source temperatures (T_H), assuming a fixed sink temperature (T_L) of TEHP is demonstrated in *Figure 14* The TEHP yields reduced COP as the value of heat source temperature goes up because for fixed heat sink temperature, as DTHP increases COP drops considerably.

Figure 15 shows the variations of heating capacity (Q_h) with I for various values of T_H assuming a fixed heat sink temperature. The system attains a maximum heating capacity of nearly 115 Watts at T_H=300K and minimum heating capacity of nearly 18Watts at T_H=375K. Hence, increase in source temperature inversely affects heating capacity of the system. The variation of COP and Q_h with I for different number of thermoelectric elements (N) is illustrated in Figure 16 and Figure 17. The heating capacity Q_h significantly goes up with number of thermoelectric elements. The maximum COP accomplished by multi-element TEHP system is higher for higher values of N.

The graphical relationships of COP and Q_h with I for different values of thermal conductance at hot and cold junction (K_H and K_L) are demonstrated in *Figure 18* and *Figure 19*. As the value of K_H and K_L increases, maximum value attained by COP decreases and Q_h increases ultimately.

Figure 20 and Figure 21 show the relationship of Q_h with current for different values of N and thermal conductance (K_H and K_L) with addition of an external irreversibility K_0 in the system. This external irreversibility K_0 is because of this additional heat leak between source/sink, due to which there is considerable drop in the value of maximum heating capacity attained by the system, as we compare these values from Figure 17 and Figure 19. Correspondingly, by comparing Figures 22 to 23 and Figures 16 to18, one can make out the fall in COP of TEHP due to heat leak between source and sink.

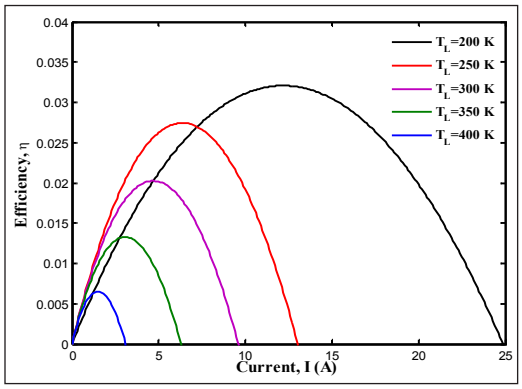


Figure 11. Thermal efficiency vs 'I' with different values of 'T_L' and fixed 'T_H' for TEG with heat leak between source and sink

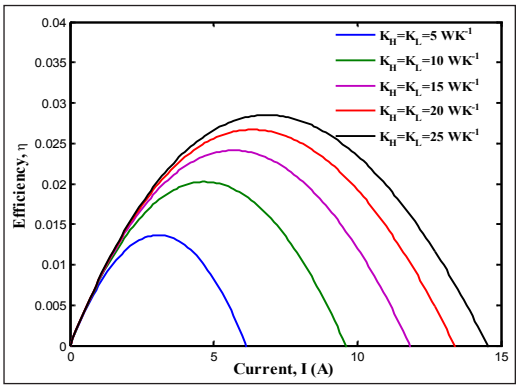


Figure 12. Thermal efficiency vs 'I' with different values of K_H and K_L for TEG with heat leak between source and sink

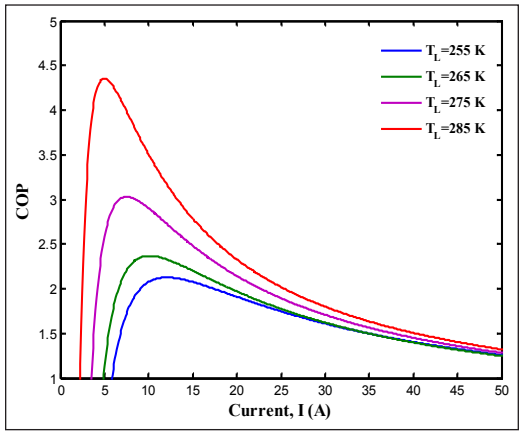


Figure 13. COP vs 'I' with different values of T_L and fixed value of T_H for TEHP

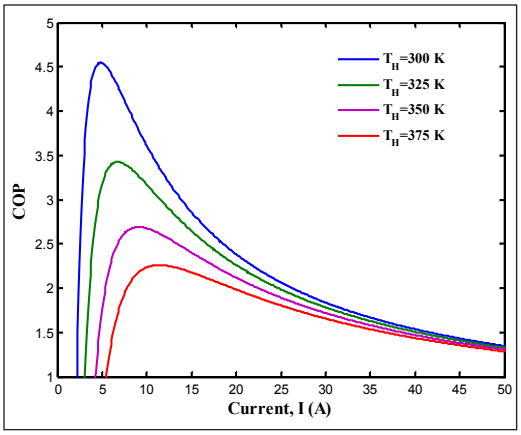


Figure 14. COP vs 'I' with different values of T_H and fixed value of T_L for TEHP

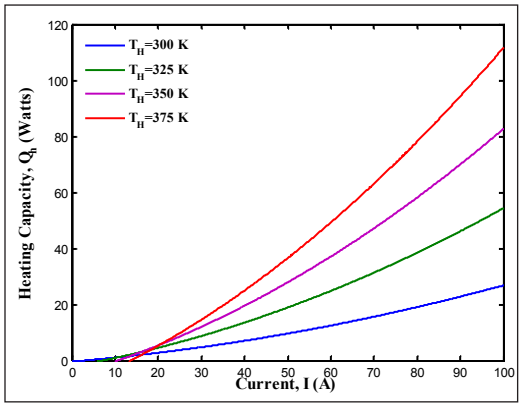


Figure 15. Heating capacity vs 'I' with different values of T_H and fixed value of T_L for TEHP

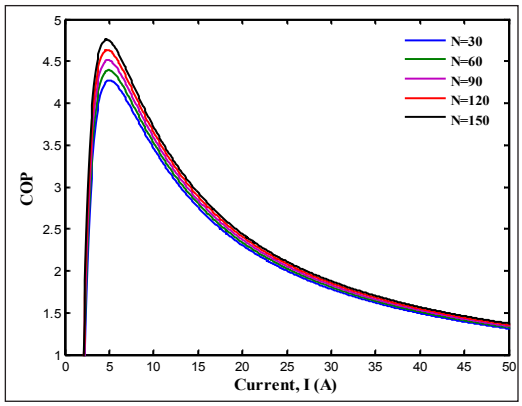


Figure 16. COP vs 'I' with different values of number of thermoelectric elements for TEHP

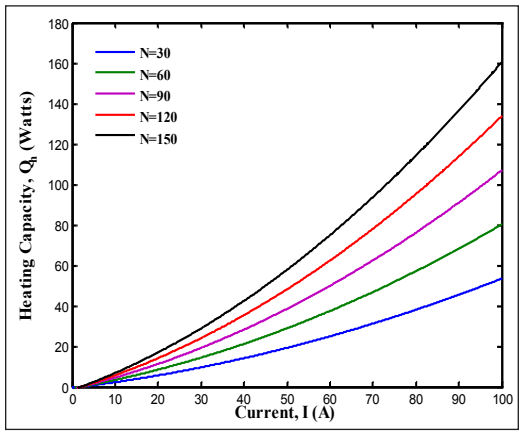


Figure 17. Heating capacity vs 'I' with different values of number of thermoelectric elements for TEHP

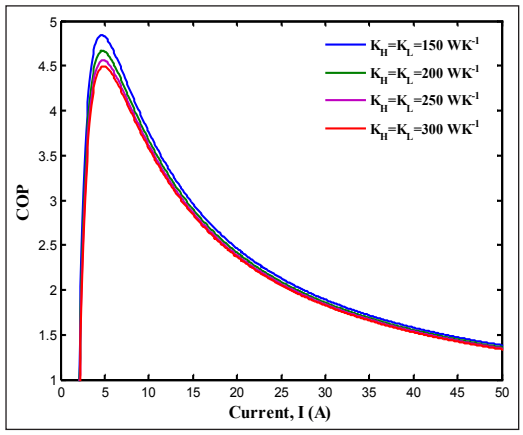


Figure 18. COP vs 'I' with different values of K_H and K_L for TEHP

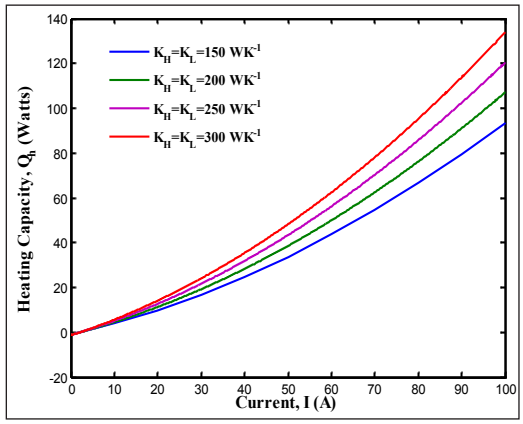


Figure 19. Heating capacity vs 'I' with different values of K_H and K_L for TEHP

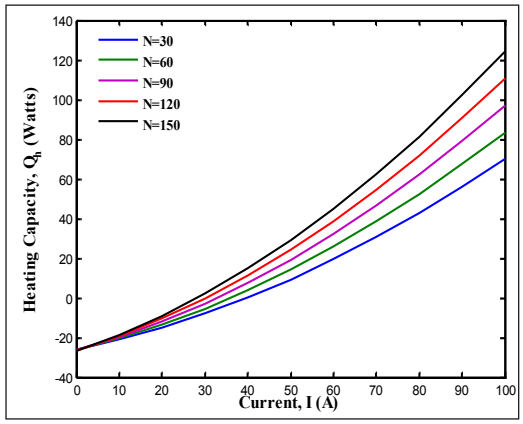


Figure 20. Heating capacity vs 'I' with different values of 'N' for TEHP with heat leak between source and sink

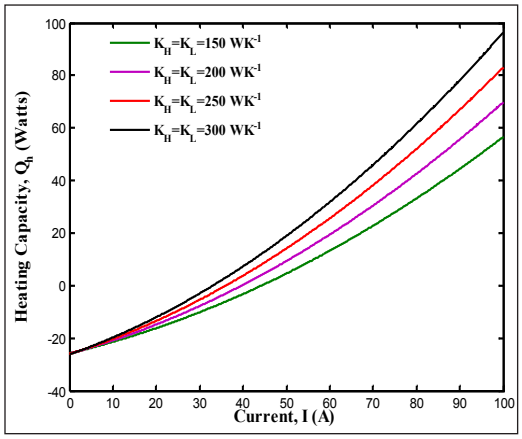


Figure 21. Heating capacity vs 'I' with different values of K_H and K_L for TEHP with heat leak between source and sink

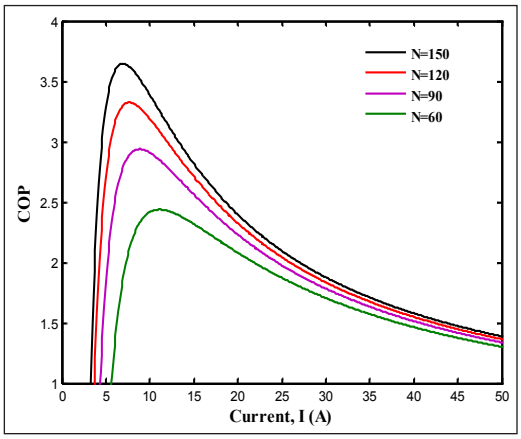


Figure 22. COP vs 'I' with different values of 'N' for TEHP with heat leak between source and sink

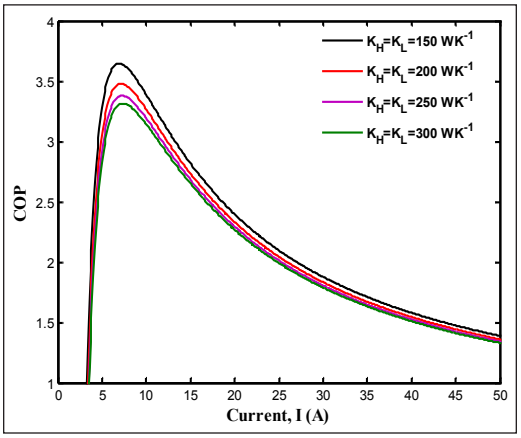


Figure 23. COP vs 'I' with different values of K_H and K_L for TEHP with heat leak between source and sink

Effect of Geometric Parameters on the Single-Stage TEG and the Systems

The power output/first law efficiency of TEG depends on the working electric current, I and for given design parameters, there exists optimal electric current related to which peak power and thermal efficiency can be obtained. Similarly, maximum COP and heating capacity can be obtained for optimal values of working electrical current. From the previous literature, it can be observed that geometry of thermoelectric elements can significantly affect the performance parameters viz. output power, first law efficiency, COP and heating capacity of thermoelectric devices. The geometric parameters comprises length and area of thermoelectric elements.

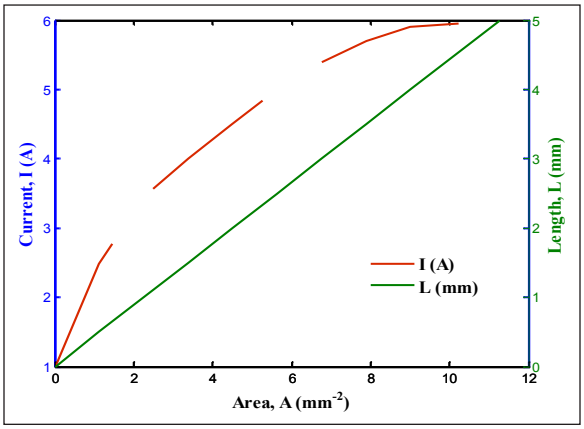


Figure 24. Variations of current and length with area of the thermoelectric element

Figure 24 shows the variations of electric current and length with area of thermoelectric couples. It is observed that length possess a linear relationship with the corresponding area of thermoelectric elements whereas the working electrical current initially increases with increase in area and stabilizes afterwards. It is clearly shown that the working electrical current is influenced by the geometric parameters of thermoelectric elements. Therefore, variations in length and area can considerably vary the required performance characteristics of thermoelectric devices. These geometric parameters are associated with the economical aspects of designing and hence, their optimum value should be chosen by appropriate numerical computations.

Figure 25 demonstrates the variations of power output with area for different values source side temperature T_H for an irreversible single stage TEG system. For numerical computation the values of other parameters are chosen as T_L=300K, K_H=K_L=10W/K, L=0.5mm, I=5A and N=127.For a fixed value of sink side temperature T_L, as T_H goes up the value of $\Delta T(T_H-T_L)$ increases and as a result the corresponding power output enhances. Initially the power output sharply rises with area 'A' but subsequently attains almost steady values. As the area of thermoelectric element increases more thermoelectric material is required which directly affects the cost of the system.

The parallel variations are observed for the thermal efficiency of an irreversible single stage TEG system with area, for different values of T_H in Figure 26. For numerical computation the values of other parameters are chosen as T_L=300K, K_H=K_L=10W/K, L=0.5mm, I=5A and N=127.For a fixed value of T_L as T_H enhances, higher values of thermal efficiency are attained for higher ΔT (T_H-T_L). The thermal efficiency increases with the area of thermoelectric elements initially whereas, attains almost steady values afterwards. It is recommended to choose the optimum values of area of thermoelectric elements from economical point of view.

Figure 27 and Figure 28 demonstrate the effects of area of thermoelectric couples on output power and thermal efficiency for a predetermined value of T_H and various values of T_L. For numerical computation the values of other parameters are chosen as T_H=450K, K_H=K_L=10W/K, L=0.5mm, I=5A and N=127.As the value of ΔT (T_H-T_L) rises, both power output and thermal efficiency consequently goes up. Hence, the temperature difference source and sink side plays an important in obtaining higher values of performance characteristics of thermoelectric devices. In the beginning, there exists a linear relationship of area of thermoelectric elements and performance parameters, but afterwards power output and efficiency get stabilize.

Figure 29 and Figure 30 exhibit the effects of area of thermoelectric couples on output power and thermal efficiency for different values of external thermal conductance. For numerical computation the values of other parameters are chosen as T_H=450K, T_L=300K, L=0.5mm, I=5A and N=127. The output power increases gradually as external thermal conductance increases.

Figure 31 and Figure 32 demonstrate the effects of area of thermoelectric couples on coefficient of performance (COP) and heating capacity for some predetermined values of T_L and various T_H values. For numerical computation the values of other parameters are chosen as T_L=250K, K_H=K_L=100W/K, L=0.5mm, I=9A and N=127.As the value of ΔT (T_H-T_L) rises, COP of an irreversible single stage TEHP system consequently goes down. With the increase in area, COP initially increases and attains a stable value afterwards. Hence, to design a practical space conditioning system one should judiciously chose the area of thermoelectric elements as area effect the economic aspects of the system. *Figure 32* illustrates the relationship of area and heating capacity for an irreversible single stage TEHP system for a predetermined value of T_L and different values of T_H. The heating capacity possesses inverse relationship with area and ΔT (T_H-T_L).

Figure 33 and Figure 34 exhibit the effects of area of thermoelectric couples on coefficient of performance (COP) and heating capacity for a predetermined T_H value and different T_L values. For numerical computation the values of other parameters are chosen as T_H=300K, K_H=K_L=100W/K, L=0.5mm, I=9A and N=127.As the value of ΔT (T_H-T_L) rises, COP of an irreversible single stage TEHP system subsequently declines. Figure 34 illustrates the relationship of area and heating capacity for an irreversible single stage TEHP system for a predetermined value of T_H and different values of T_L. The heating capacity possesses inverse relationship with area and ΔT (T_H-T_L).

Figure 35 and Figure 36 exhibit the effects of area of thermoelectric couples on COP and heating capacity for different values of external thermal conductance. For numerical computation the values of other parameters are chosen as T_H=300K, T_L=250K, L=0.5mm, I=9A and N=127. The COP of an irreversible single stage TEHP system declines with rise in external thermal conductances. The heating capacity of the system declines with the rise

in area of thermoelectric couples as shown in Figure 36. Hence, COP and heating capacity attain their maximum values for different values of area of thermoelectric elements and for practical design one has to optimize both simultaneously.

The similar effect has been observed for the length of thermoelectric couples on output power /thermal efficiency for TEG system and COP/heating capacity for TEHP system. The direct relationship of area and length has been shown in Figure 24. It is concluded from above studies that for geometric parameters put considerable effect on the performance parameters as well as economic aspects of designing of thermoelectric devices. Cost of thermoelectric material is concerned with the geometry of thermoelectric elements. For real designing, performance and cost should be optimized simultaneously in order to achieve most effective operation of thermoelectric devices in minimum possible cost.

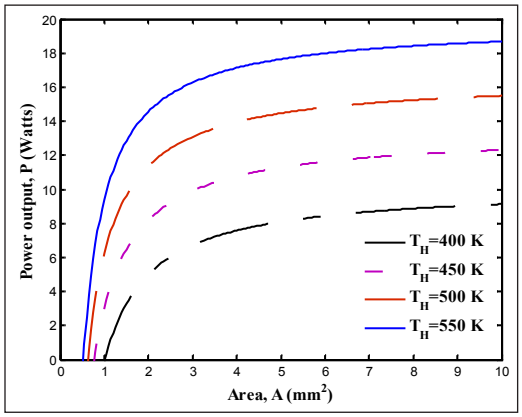


Figure 25. Power output vs area with different values of 'T_H' and fixed value of T_L for TEG

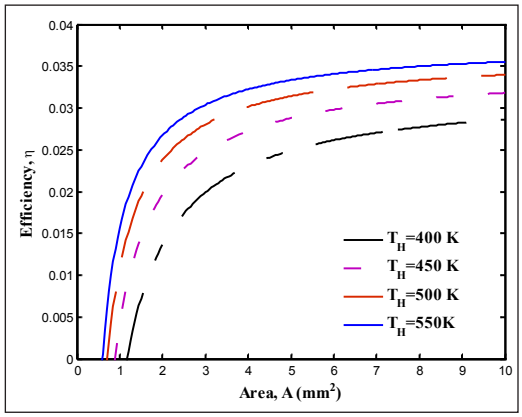


Figure 26. Efficiency vs area with different values of 'T_H' and fixed value of T_L for TEG

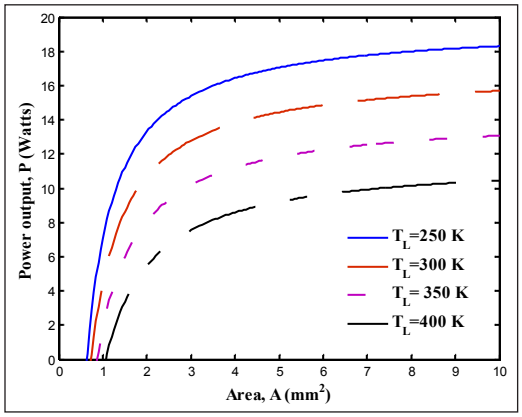


Figure 27. Power output vs area with different values of 'T_L' and fixed value of T_H for TEG

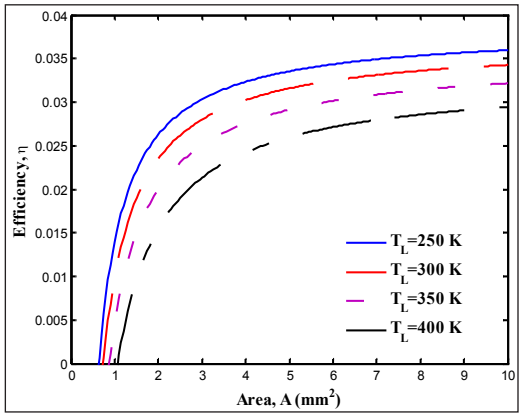


Figure 28. Thermal efficiency vs area with different values of 'T_L' and fixed value of T_H for TEG

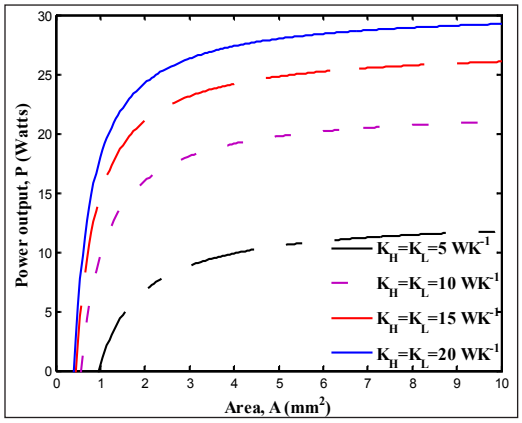


Figure 29. Power output vs area with different values of 'K_H' and 'K_L' for TEG

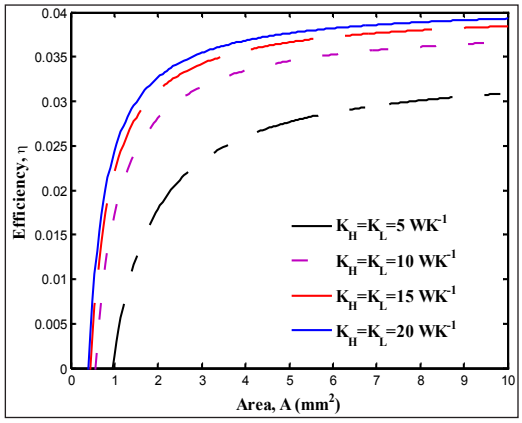


Figure 30. Thermal efficiency vs area with different values of 'K_H' and 'K_L' for TEG

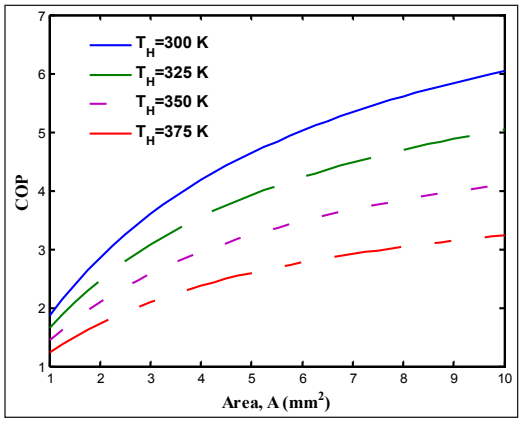


Figure 31. COP vs area with different values of 'T_H' and fixed value of 'T_L' for TEHP

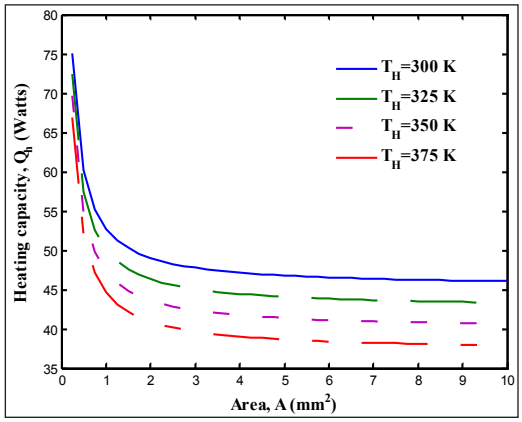


Figure 32. Heating capacity vs area with different values of 'T_H' and fixed value of 'T_L' for TEHP

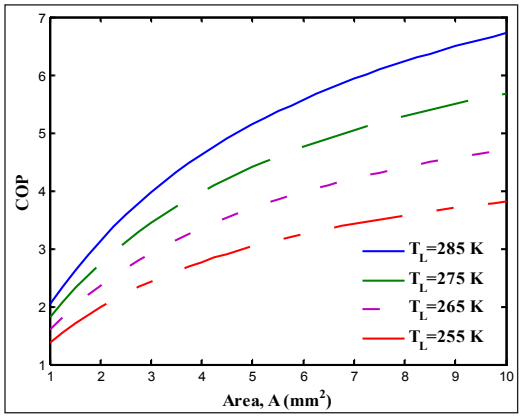


Figure 33. COP vs area with different values of 'T_L' and fixed value of 'T_H' for TEHP

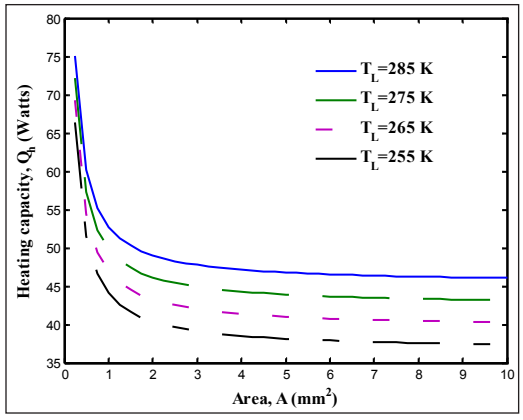


Figure 34. Heating capacity vs area with different values of 'T_L' and fixed value of 'T_H' for TEHP

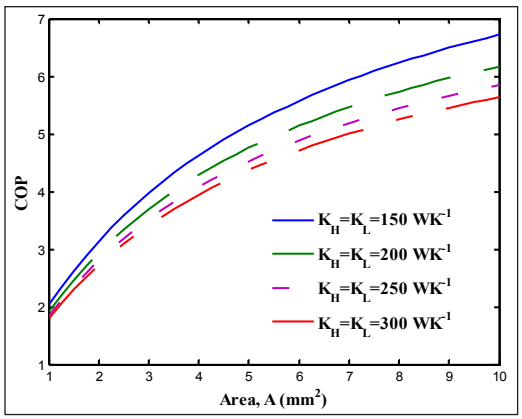


Figure 35. COP vs area with different values of 'K_H' and 'K_L' for TEHP

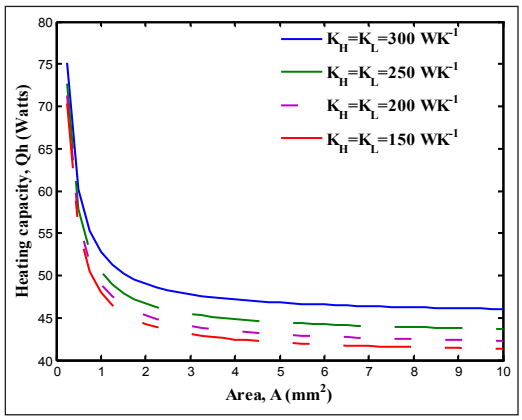


Figure 36. Heating capacity vs area with different values of 'K_H' and 'K_L' for TEHP

Validation of the Study

In order to verify/validate the thermodynamic modeling of TEG and TEHP systems, the data stated in the previous studies are utilized and the comparison is done as illustrated in Table 3. For TEG, the obtained outcomes show an agreement with that of stated in Hogblom and Andersson (2014). Conversely, COP of TEHP obtained in the present work matches with that of stated in Nami et al. (2017) whereas heating capacity matches with that of reported in Kuashik et al. (2015), on the basis of same operating conditions and input parameters.

Table 3
Comparison of designing and performance parameters with previous literature

Comparative Investigations	Designing Parameters								Performance Parameters			
	TEG				TEHP				TEG		TEHP	
	I (A)	T_h (K)	T_c (K)	N	I (A)	T_h (K)	T_c (K)	N	P (W)	η	Q_H **(W)**	**COP**
Present Study	2.5	400	300	127	5.12	300	250	127	4.85	0.02	15.12	4.6
Hogblom and Andersson (2014)	4.26	158	65	-	-	-	-	-	5.7	-	-	-
Kaushik et al. (2015)	-	-	-	-	5.65	313	303	-	-	-	14.72	-
Nami et al. (2017)	-	-	-	-	5.6	310	300	-	-	-	-	3.76

CONCLUSION

Finite time thermodynamic analysis carried out on multi-element irreversible TEG and TEHP devices directs the major concluding points as follows:

- The generalized finite time thermodynamic models of TEG and TEHP systems are established considering inner geometrical dimensions, heat transfer in external mode and temperature dependent physical properties.
- Formulation of performance parameters of thermoelectric devices in context with electric current, thermoelectric couples, temperature of hot/cold side and thermal conductances of hot/cold side is presented.
- The analytical formulation of hot/cold sides temperature based on energy balanced equation is achieved for TEG and TEHP devices. The proposed system configurations are investigated with and without heat leak between heat source and sink.
- The effects of design factors on power output/first law efficiency and heating capacity/ coefficient of performance are analyzed.
- The effect of area and length on power output/thermal efficiency, COP and heating capacity is investigated with variations in T_H, T_L, K_H/K_L for single stage TEG and TEHP systems.
- It is concluded from above studies that for geometrical parameters put considerable effect on the performance parameters as well as economic aspects of designing of thermoelectric devices. Cost of thermoelectric material is concerned with the geometry of thermoelectric elements. For real designing, performance and cost should be optimized simultaneously in order to achieve most effective operation of thermoelectric devices.
- It is also observed that various parameters impact the design variables in different ways and performance parameters of thermoelectric devices. The thermodynamic optimization is mandatory and operative for practical operations at variable conditions. The varying features of various input and geometric/design variables obtained in this study is utilized as the benchmark for assessing the effects of external/internal irreversibility on optimum performances of TEG and TEHP systems. Moreover, a practical setup of the proposed thermoelectric systems could be fabricated based on the obtained optimal values of design parameters.

REFERENCES

Andresen, B. (1983). *Finite-time Thermodynamics*. Physics Laboratory II, University of Copenhagen.

Angrist, S. W. (1992). *Direct Energy Conversion*. Boston: Allyn and Bacon Inc.

Chen, J., & Wu, C. (2000). Analysis of the performance of a thermoelectric generator. *Journal of Energy Resource-ASME, 122*(2), 61-63.

Chen, L., Li, J., Sun, F., & Wu, C. (2005a). Performance optimization of two-stage semiconductor thermoelectric-generator. *Applied Energy, 82*(4), 300-312.

Chen, L., Li, J., Sun, F., & Wu, C. (2005b). Effect of heat transfer on the performance of two stage semiconductor thermoelectric refrigerators. *Journal of Applied Physics, 98*(3), 1-7. Chen, L., Li, J., Sun, F., & Wu, C. (2008). Performance optimization for a two-stage thermoelectric heat-pump with internal and external irreversibilities. *Applied Energy, 85*(7), 641-49.

Chen, L., Sun, F., & Chen, W. (1994). Finite time thermodynamic analysis for a thermoelectric refrigerator and heat pump. *Chinese Journal of Engineering Thermophysics*, *15*(1), 13–16.

Chen, W. H., Liao, C. Y., & Hung, C. I. (2012). A numerical study on the performance of miniature thermoelectric cooler affected by Thomson effect. *Applied Energy, 89*(1), 464-473.

David, B., Ramousse, J., & Luo, L. (2012). Optimization of thermoelectric heat pumps by operating condition management and heat exchanger design. *Energy Conversion and Management, 60*, 125-133.

Disalvo, F. J. (1999). Thermoelectric cooling and power generation. *Science, 285*(5428), 703- 706.

Gaurav, K., & Pandey, S. K. (2017). Efficiency calculation of a thermoelectric generator for investigating the applicability of various thermoelectric materials. *Journal of Renewable and Sustainable Energy, 9*(1), 014701.

Goldsmith, H. J. (1964). *Thermoelectric Refrigeration* (2nd ed.). New York: Wiley.

Hans, R., Kaushik, S. C., & Manikandan, S. (2016). Performance optimization of two-stage exoreversible thermoelectric heat pump in electrically series, parallel and isolated modes. *International Journal of Energy Technology and Policy, 12*(4), 313-332.

Hans, R., Manikandan, S., & Kaushik, S. C. (2015). Performance optimization of two-stage exoreversible thermoelectric converter in electrically series and parallel configuration. *Journal of Electronic Materials, 44*(10), 3571-3580.

Hogblom, O., & Andersson, R. (2014). Analysis of Thermoelectric Generator Performance by Use of Simulations and Experiments. *Journal of Electronic Materials, 43*(6), 2247-2254.

Kaushik, S. C., Manikandan, S., & Hans, R. (2015). Energy and exergy analyses of thermoelectric heat pump systems. *International Journal of Heat and Mass Transfer, 86*, 843-852.

Kim, Y. W., Ramousse, J., Fraisse, G., Dalicieux, P., & Baranek, P. (2014). Optimal sizing of a thermoelectric heat pump for heating energy efficient buildings. *Energy Buildings, 28*(70), 106-116.

Nami, H., Nemati, A., Yari, M., & Ranjbar, F. (2017). A comprehensive thermodynamic and exergoeconomic comparison between single- and two-stage thermoelectric cooler and heater. *Applied Thermal Engineering, 124*, 756-766.

Riffat, S. B., & Ma, X. (2003). Thermoelectric: a review of present and potential applications. *Applied Thermal Engineering, 23*(8), 913-35.

Riffat, S. B., & Ma, X. (2004). Improving the coefficient of performance of thermoelectric cooling systems: a review. *International Journal of Energy Research, 28*(9), 753-768.

Rowe, D. M. (2005). *CRC Handbook of Thermoelectrics*. Boca Raton, Florida: CRC press.

Sun, X., Liang, X., Shu, G., Tian, H., Wei, H., & Wang, X. (2014). Comparison of the two-stage and traditional single-stage thermoelectric generator in recovering the waste heat of the high temperature exhaust gas of internal combustion engine. *Energy, 77*, 489-498.

Tan, H., Fu, H., & Yu, J. (2017). Evaluating optimal cooling temperature of single-stage thermoelectric cooler using thermodynamic second law. *Applied Thermal Engineering, 123*, 845-851.

Xuan, X. C., Ng, K. C., Yap, C., & Chua, H. T. (2002). The maximum temperature differences and polar characteristics of two-stage thermoelectric coolers. *Cryogenics, 42*(5), 273-278.

Zhu, J., Gao, J., Chen, M., Zhang, J., Du, Q., Rosendahl, L. A., & Suzuki, R. O. (2011). Experimental study of a thermoelectric generation system. *Journal of Electronic Materials, 40*(5), 744-752.

Pertanika J. Sci. & Technol. 26 (4): 1999 - 2008 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Homemade Pectin, Made of *Musa Acuminata Cavendish*'s (Cavendish Banana) and *Musa Paradisiaca Formatypica*'s (Horn Shaped Banana) Skin, and Its Effectiveness Analysis as Heavy Metal Absorber by Using Laser-Induced Breakdown Spectroscopy (LIBS)

Nadia Adiasa[1,2], Lidya Kartika[2], Koo Hendrik Kurniawan[1], Zener Sukra Lie[3] and Rinda Hedwig[3]*

[1]Maju Makmur Mandiri Research Center, 40 Srengseng Raya, Jakarta 11630, Indonesia
[2]Tunas Manunggal Foundation, Sekolah Santa Laurensia, 1 Sutera Utama, Alam Sutera, Serpong, Banten 15320, Indonesia
[3]Computer Engineering Department, Faculty of Engineering, Bina Nusantara University, 9 K.H. Syahdan, Jakarta 11480, Indonesia

ABSTRACT

Heavy metal pollution in water, especially in the North Jakarta area, has become severe from time to time. A heavy metal adsorbent is one of the alternative solutions to clean the waste water from industries before it is released into the environment. In this paper, we create pectin that has been proven to be an effective absorbent for heavy metal (Pb) and is made of *Musa* ac*uminata Cavendish* (cavendish banana) and *Musa paradisiaca Formatypica* (horn shaped banana). These two banana species can be easily found in the markets of Jakarta. Therefore, we compare the two species to find out which is more effective to use as heavy metal adsorbent. The selection is based on the reduction of the Pb concentration after the introduction of banana pectin by means of Pb emission intensity reading taken using commercially available laser-induced breakdown spectroscopy (LIBS) system. The effectiveness of pectin from the two banana skins as Pb adsorbent is comparable either in pellet form or solution in water. The result shows that the average effectiveness of homemade pectin as Pb adsorbent is 37.5%, which is lower than the commercial pectin available in the market.

ARTICLE INFO

Article history:
Received: 28 February 2018
Accepted: 10 July 2018
Published: October 2018

E-mail addresses:
saebaryo12324@gmail.com (Nadia Adiasa)
l_dyakartika@yahoo.com (Lidya Kartika)
kurnia18@cbn.net.id (Koo Hendrik Kurniawan)
zener_lie@yahoo.com (Zener Sukra Lie)
rinda@binus.edu (Rinda Hedwig)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

In the future, the technique of producing homemade pectin will be improved to get better Pb absorption efficiency.

Keywords: Banana, heavy metal, LIBS, Pb adsorbent, pectin

INTRODUCTION

Banana, with more than 50 varieties, is one of the main fruits well grown in tropical countries such as Indonesia (Setyobudi, 2000). It has nutritional value and composes 45% of the fruits consumed by most people (Hapsari & Lestari, 2016; Setiawan, Sulaeman, Giraud, Driskell, 2011). Since many of Indonesia's cuisine, especially snacks, use bananas as their basic ingredient, there are a lot of banana skins left as waste. The waste, in fact, has been used for biomass fuel production, organic fertilizer, and organic cattle food in the farm. However, it is also interesting to use it for other useful things.

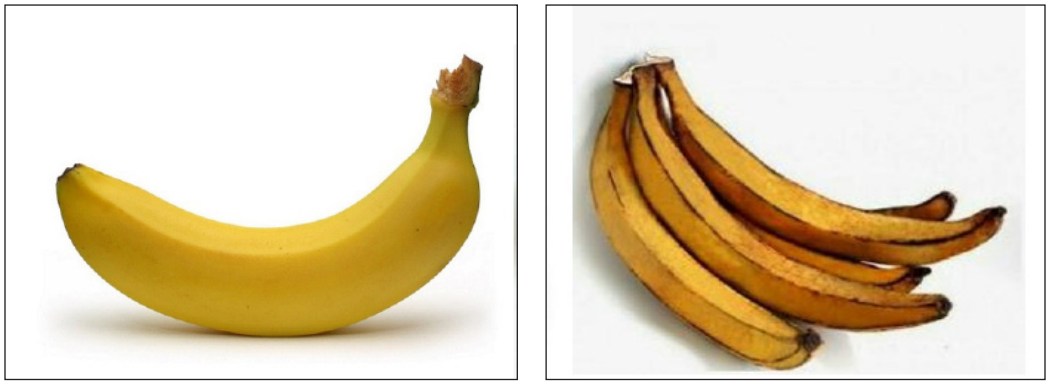
In this paper, we introduce how to recycle the banana skins into pectin (Emaga, Ronkart, Robert, Wathelet, & Paquot, 2008; Mohnen, 2008) in the traditional way. Pectin is widely used in the industry as gelling agent and stabilizer. It is used in jams, jellies, frozen foods, low-calorie foods, or as fat/sugar replacer (Thakur, Singh, Handa, & Rao, 1997) and at the same time, the pectin can also be used as a heavy metal adsorbent in water (Li, Yang, & Zhao, 2007; Wan Ngah, & Hanafiah, 2008).

Heavy metal pollution in water especially in the coastal areas of Indonesia is severe and is mainly caused by the waste of industries. (Hosono, Su, Delinom, Toyotae, Kaneko, & Taniguchi, 2011; Nienhuis, 1986). Pectin can be one of the alternate solutions for reducing the heavy metal waste in water, and for this purpose, we make pectin that can be used as a low-cost adsorbent of heavy metals waste from industries. The pectin is mixed with the waste water that is able to transform into gel structure, which leaves the uncontaminated waste water as a result. In fact, pectin can be produced from any kind of fruit skin. However, in this paper, we focus on banana skins which are easily found in Indonesia. Finally, the adsorption effectiveness of banana pectin made in this study is applied as a Pb absorbent in waste water.

MATERIALS AND METHODS

Preparation of Banana Pectin

In this experiment, we used the skin of banana from genus *Musa Acuminata Cavendish* (Cavendish banana) and *Musa Paradisiaca Formatypica* (horn shaped banana) as the basic material of pectin, that is utilized as a low-cost heavy metals adsorbent, as shown in figure 1. The skin structure of both bananas is different in that the skin of Cavendish banana is softer than horn shaped one. Both skins were 4 cm diced and heated in the 500 watts oven until they dried separately with water concentration of less than 5%. Later, the dried skin



(a) (b)

Figure 1(a). Musa Acuminata Cavendish (cavendish banana), and *(b). Musa Paradisiaca Formatypica* (horn shaped banana)

was blended into powder. Every 25 g of the powder and 25 g of citric acid ($C_6H_8O_7$) was dissolved in 1 liter distilled water by using stirrer. The above solution was then heated at 100°C to finally get dry pectin powder.

Further, the dried pectin powder was pulverized until its grain size was less than 50 µm. In order to check the effectiveness of homemade pectin as adsorbent of Pb heavy metals in water, we prepared four 50 ml beaker glasses, each of which is filled with 40 ml water containing 0.1 g $PbCl_2$, along with: (i) no pectin, (ii) 0.1 g cavendish banana pectin, (iii) 0.1 g horn shaped banana pectin and (iv) 0.1 g commercial available pectin. The mixtures were stirred using a magnetic stirrer for 30 minutes at room temperature, and was then filtered through 20 µm filter paper. The resulting solution was used as samples. A separate experiment was carried out to check the absorption rate of Pb by the pectin, using the same technique as mentioned above, utilizing only cavendish banana pectin. The mixtures are stirred using a magnetic stirrer for between 5 to 60 minutes at room temperature, at 5 minute increments, and was then filtered through 20 µm filter paper. The resulting solution was used as samples.

Laser-Induced Breakdown Spectroscopy Measurement

The schematic diagram of the experimental setup for spectroscopic measurement is given by Figure 2 as we have reported previously (Suyanto et al., 2012). The laser used in this experiment was a nano-second (ns) Nd:YAG laser (New Wave, 200 mJ maximum energy, 1064 nm, 8 ns) operated in Q switched mode at 10 Hz repetition rate and a reduced output energy of 100 mJ. The laser beam was directed onto the sample in the chamber through a quartz window of the chamber. A convex lens of 150 mm focal length was used with -5 mm defocused condition to yield a power density of around 250 MW/cm^2 on the sample surface.

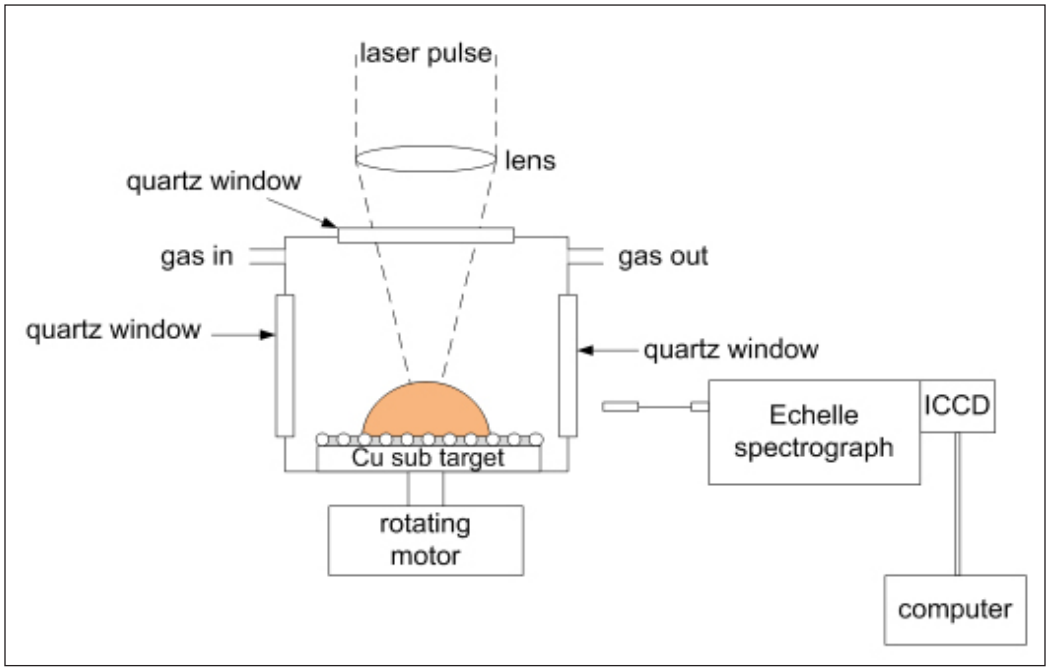


Figure 2. Experimental setup

The choices of the laser energy and focusing condition were determined in another preliminary experiment searching for the minimum ablation power density needed to yield stable and reproducible analyte intensities from the powder sample without inducing the Cu subtarget emission lines from the sample holder and causing blowing-off of the sample powder in a low pressure ambient gas. The chamber was equipped with an inlet and outlet ports for the continuous flow of ambient gas at a constant flow-rate to maintain the gas pressure at 30 Torr. The chamber had an additional quartz window for the observation of plasma emission. The plasma emission intensity was collected at one end of an optical fiber of 10 μm core diameter, with its other end connected to the entrance slit of an Echelle spectrograph (Mechelle M500 type equipped with a gated ICCD system, Andor iStar ICCD, 200-975 nm wavelength) for the measurement of emission spectrum in a broad spectral range. The samples were rotated during the measurement so that each laser pulse would hit a different spot on the sample surface.

RESULTS AND DISCUSSION

In order to know the rate of Pb absorption after the introduction of the pectin, the prepared samples as described in the previous section is used. Figure 3 shows Pb absorption rate after the introduction of cavendish banana pectin in a mixture of 40 ml water and 0.1 g $PbCl_2$. The emission intensity of Pb I 405.7 nm was obtained by irradiating the sample

surface using 100 mJ laser energy in 30 Torr ambient air. The ICCD system was set at time integrated mode. One clearly saw that the maximum absorption of Pb element by the cavendish pectin occurs at around 30 minutes and almost no increment of Pb absorption even though the stirring time is increased up to 60 minutes. Based on this result, for the next experiment, 30 minutes stirring time was used. The spectrum of cavendish banana pectin and horn shaped banana pectin can be seen in Figure. 4(a) for cavendish banana and (b) for horn shaped banana.

In order to check the effectiveness of homemade pectin, four samples prepared as mentioned in the experimental procedures is used and the result is presented in Fig. 5 (a) without pectin; (b) with 0.1 g cavendish banana pectin; (c) with 0.1 g horn-shaped banana pectin and (d) with 0.1 g commercial pectin. By comparing Figure 5(a) with Figures 5(b) and 5(c), one sees that the maximum Pb absorption is almost the same (30%) for both banana type. Meanwhile total removal of Pb as shown in Figure. 5(d) is found for the case of commercial pectin.

Although the effectiveness of homemade low-cost pectin from banana skin is still lower than the commercial one, we believe that by improving the pectin preparation, through smaller pectin grain size as well as controlling the liquid temperature, the effectiveness of home made pectin will be enhanced. Further experiments are required and pectin should also be tested for other heavy metals.

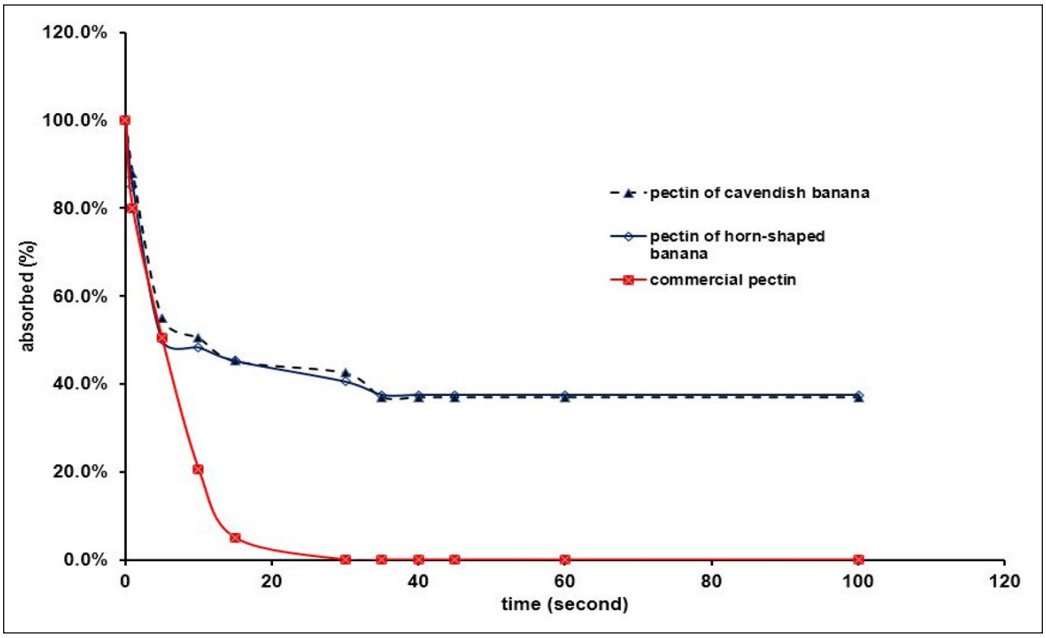


Figure 3. Pb I 405.7 nm emission intensity as a function of time after pectin introduction into the water which is containing 0.1 g $PbCl_2$.

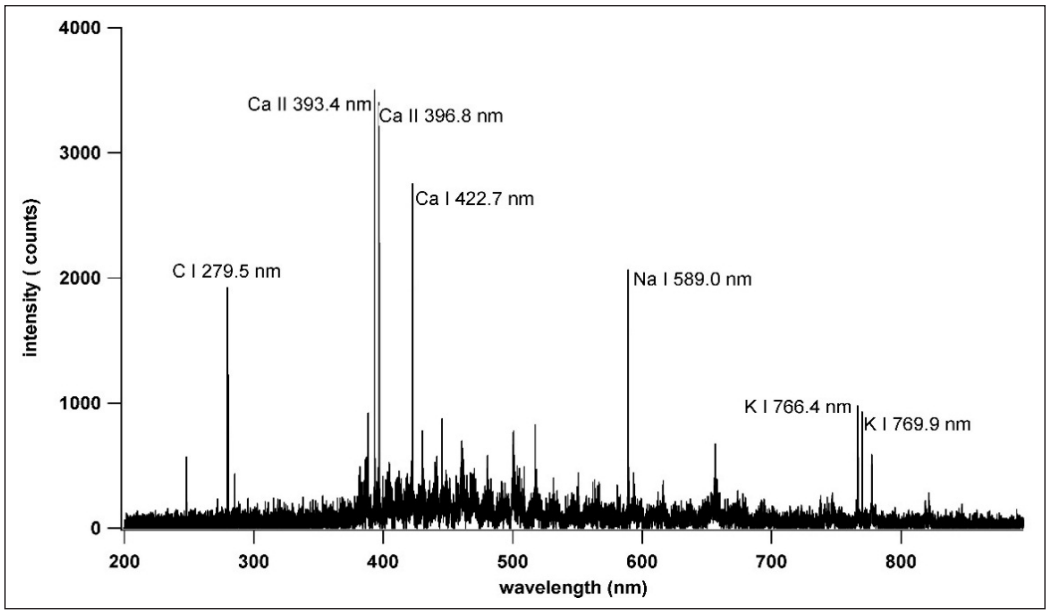


Figure 4(a). Spectrum of pectin made of cavendish banana.

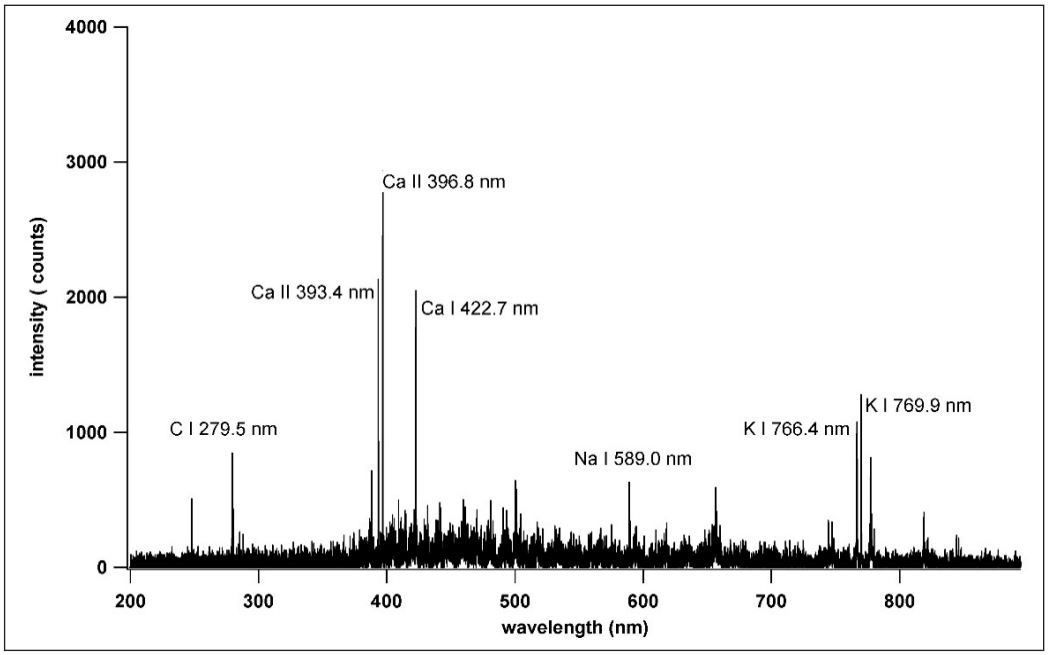


Figure 4(b). Spectrum of pectin made of horn shaped banana.

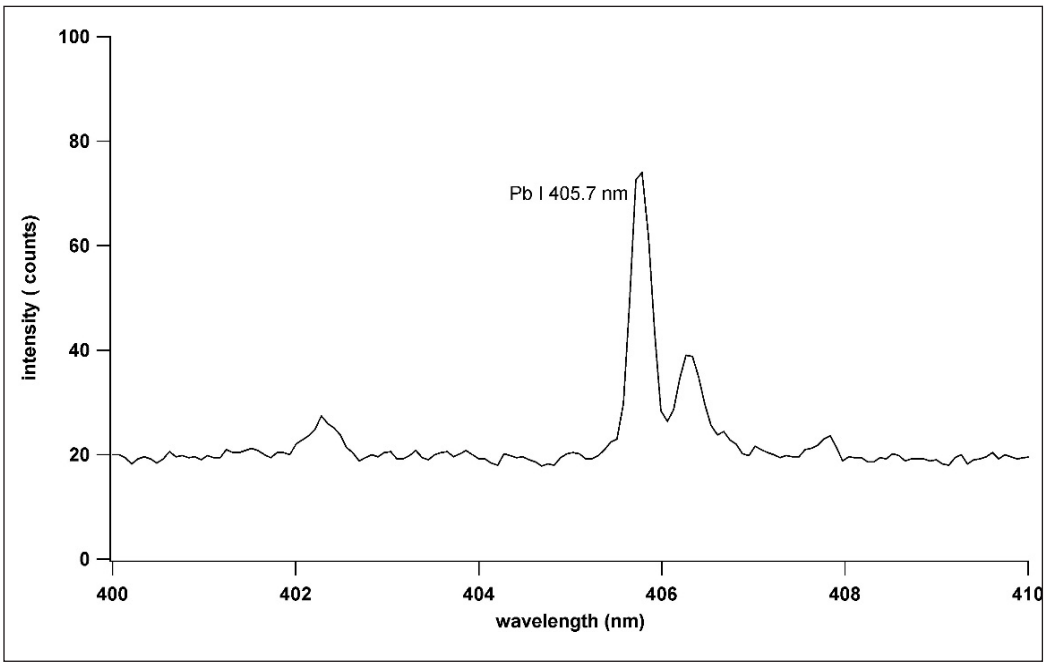


Figure 5(a). The effectiveness of Pb I 405.7 nm adsorption in liquid by using pectin and the spectrum of 40 ml H_2O + 0.1 g $PbCl_2$

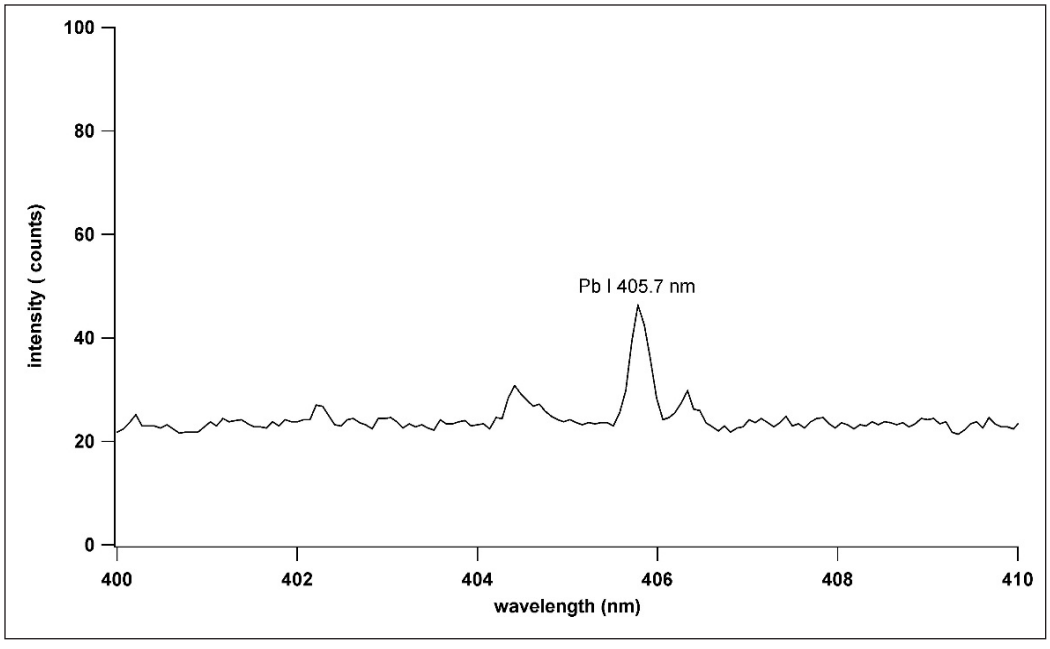


Figure 5(b). 40 ml H_2O + 0.1 g $PbCl_2$ + 0.1g pectin of cavendish banana

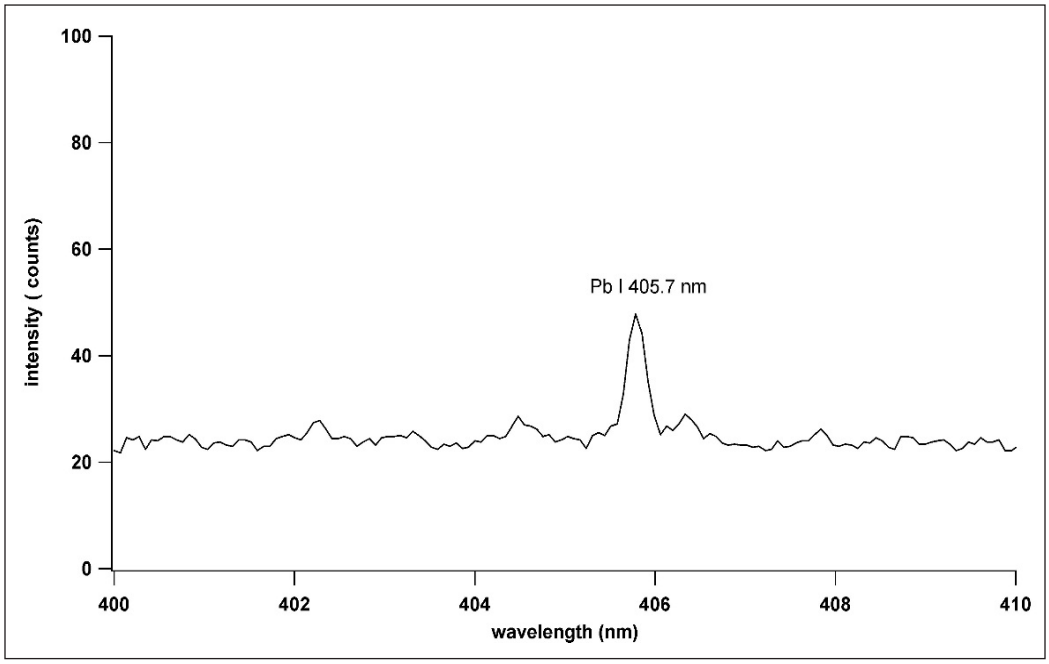


Figure 5(c). 40 ml H_2O + 0.1gr $PbCl_2$ + 0.1 g pectin of horn shaped banana

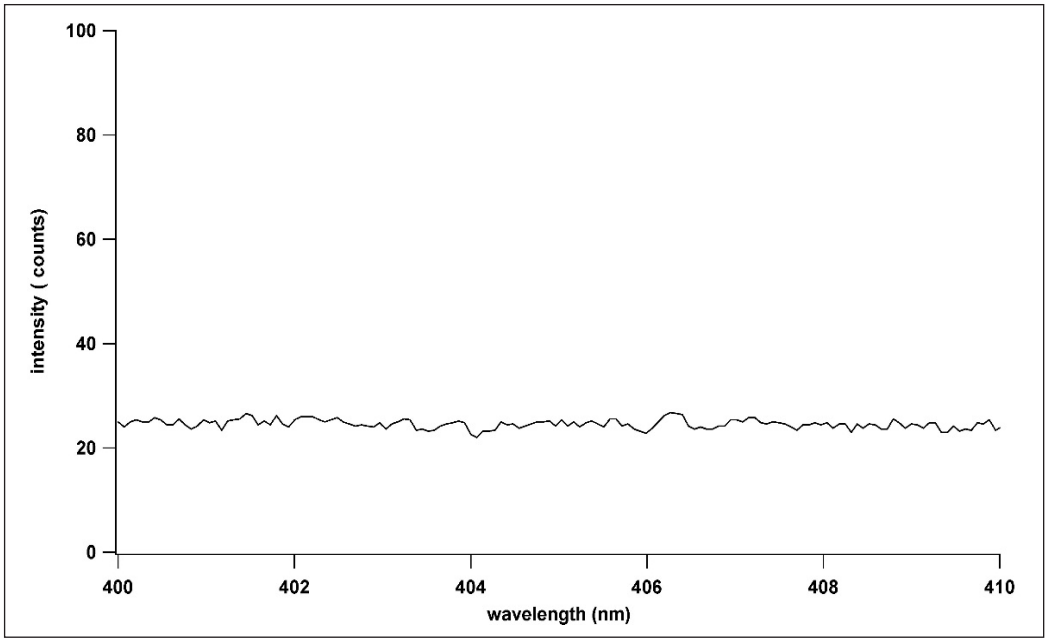


Figure 5(d). 40 ml H_2O + 0.1gr $PbCl_2$ + 0.1g commercial pectin.

CONCLUSION

Pectin is a useful compound that is used on a daily basis in the pharmaceutical or food and beverages industries. Moreover, pectin can be used as an alternate solution to adsorb heavy metals elements in waste water. Although many kinds of fruit can be used as pectin, each will produce the pectin differently, and the selection of banana skin in this experiment proves that their waste can be used as pectin. Although their effectiveness as adsorbent is only 37.5%, we still can improve the way to make the low-cost pectin by improving the sample preparation in the future.

REFERENCES

Emaga, T. H., Ronkart, S. N., Robert, C., Wathelet, B., & Paquot, M. (2008). Characterisation of pectins extracted from banana peels (Musa AAA) under different conditions using an experimental design. *Food Chemistry, 108*(2), 463-471.

Hapsari, L., & Lestari, D. A. (2016). Fruit characteristic and nutrient values of four indonesian banana cultivars (Musa spp.) At different genomic groups. *Agrivita, 38*(3), 303-311.

Hedwig, R., Budi, W. S., Abdulmadjid, S. N., Pardede, M., Suliyanti, M. M., Lie, T. J., … & Tjia, M. O. (2006). Film analysis employing subtarget effect using 355 nm Nd-YAG laser-induced plasma at low pressure. *Spectrochimica Acta - Part B Atomic Spectroscopy, 61*(12), 1285-1293.

Hosono, T., Su, C. C., Delinom, R., Umezawa, Y., Toyotae, T., Kaneko, S., & Taniguchi, M. (2011). Decline in heavy metal contamination in marine sediments in Jakarta Bay, Indonesia due to increasing environmental regulations. *Estuarine, Coastal and Shelf Science, 92*(2), 297-306.

Kagawa, K., Lie, T. J., Hedwig, R., Abdulmajid, S. N., Suliyanti, M. M., & Kurniawan, K. H. (2000). Subtarget effect on laser plasma generated by transversely excited atmospheric CO2 laser at atmospheric gas pressure. *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers, 39*(5R), 2643-2646.

Khumaeni, A., Budi, W. S., Wardaya, A. Y., Hedwig, R., & Kurniawan, K. H. (2016). Rapid Detection of Oil Pollution in Soil by Using Laser-Induced Breakdown Spectroscopy. *Plasma Science and Technology, 18*(12), 1186-1191.

Li, F. T., Yang, H., Zhao, Y., & Xu, R. (2007). Novel modified pectin for heavy metal adsorption. *Chinese Chemical Letters, 18*(3), 325-328.

Mohapatra, D., Mishra, S., & Sutar, N. (2010). Banana and its by-product utilisation: an overview. *Journal of Scientific and Industrial Research, 69*(5), 323-329.

Mohnen, D. (2008). Pectin structure and biosynthesis. *Current Opinion in Plant Biology, 11*(3), 266-277.

Nienhuis, P. (1986). Background levels of heavy metals in nine tropical seagrass species in Indonesia. *Marine Pollution Bulletin, 17*(11), pp. 508-511.

Setiawan, B., Sulaeman, A., Giraud, D. W., & Driskell, J. A. (2001). Carotenoid Content of Selected Indonesian Fruits. *Journal of Food Composition and Analysis, 14*(2), 169-176.

Setyobudi, L. (2000). The Indonesian banana industry. In *International symposium of bananas and food security* (pp. 227-236). Douala, Cameroon: INIBAP.

Thakur, B. R., Singh, R. K., Handa, A. K., & Rao, M. A. (1997). Chemistry and uses of pectin — A review. *Critical Reviews in Food Science and Nutrition , 37*(1), 47-73.

Wan Ngah, W. S., & Hanafiah, M. A. K. M. (2008). Removal of heavy metal ions from wastewater by chemically modified plant wastes as adsorbents: A review. *Bioresource Technology, 99*(10), 3935-3948.

Pertanika J. Sci. & Technol. 26 (4): 2009 - 2029 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Numerical Investigation of Fluid Flow and Heat Transfer in a Two-Pass Channel with Perforated Ribs

Arjumand Rasool* and Adnan Qayoum

Department of Mechanical Engineering, National Institute of Technology Srinagar, Kashmir, 190006, India

ABSTRACT

Rib turbulators are largely utilized for enhancement of heat transfer in cooling channels of gas turbine blades. The present study focuses on the heat transfer, fluid flow and pressure drop study of perforated ribs fixed to the bottom wall of a two-pass square channel. The turbulent flow details for heat transfer and fluid flow for perforated ribs are simulated by using commercial software Comsol 5.3a with an established turbulence model i.e. Standard k-ε. The assiduity is towards analyzing the possible effects of varying inclination angle (0 to 30^o) and shape of hole (cylinder to square) on heat transfer and friction factor characteristics for turbulent flow. The studied Reynolds number varies from 10000 to 50,000. Computations are carried out to determine inter-rib distribution of local heat transfer coefficient over the bottom ribbed wall. The phenomenon responsible for the heat transfer enhancement by perforated ribs is delineated. The results reveal that perforated ribs lead to enhancement of local heat transfer distribution (Nu/Nu_o) on the end wall downstream the ribs. Perforated ribs develop longitudinal vortices. These vortices cause an increase in flow mixing and turbulent kinetic energy. The square perforated ribs provide a 37.1–57.3% higher normalized average Nusselt number relative to the solid ribs, however induce high pressure drop. Overall, square perforated ribs (Case-3) provide the best thermal-hydraulic performance.

Keywords: Comsol, convective heat transfer, perforated ribs, turbine blade cooling, turbulence model

ARTICLE INFO

Article history:
Received: 1 March 2018
Accepted: 28 May 2018
Published: October 2018

E-mail addresses:
arjumand.beigh@yahoo.co.in (Arjumand Rasool)
adnan@nitsri.ac.in (Adnan Qayoum)
* Corresponding author

INTRODUCTION

Technological improvements in heat exchange processes are necessary for the thermal energy systems to combat the everyday increasing demand for power. Most of the thermal systems, particularly gas turbine blades operate under high inlet

ISSN: 0128-7680
e-ISSN: 2231-8526

temperatures (1200K-1700K) to ascertain the higher thermal efficiency. The present materials available for turbine blades are not able to withstand large periods of exposure to these harsh temperature conditions, even with thermal barrier coatings (TBC). Therefore, requires effective cooling methods. Since the introduction of turbine blade cooling, considerable improvements in increasing turbine inlet temperatures have been made. Internal passages of turbine blades are cooled by circulating air through them. These passages are connected by 180^{o} bend. Various heat transfer enhancement (active/ passive) techniques have been proposed to improve the internal cooling of thermal systems like gas turbine blade cooling, electronic systems, nuclear reactors and compact heat exchangers etc (Webb, 1994). The heat transfer enhancement and cooling techniques used in gas turbine are discussed in detail by (Han et al., 2012). For internal flows, large scale surface roughness including repeated ribs (Acharya et al., 1993; Han et al., 1985; Tariq et. al., 2003) vortex generators (Liou et al., 2000; Yuan & Tao., 2003) and baffles/ fins (Chamoli & Thakur, 2015; Gao et al., 2016) have been exhaustively employed for heat transfer augmentation. Rib turbulators on the channel walls are rigorously studied for heat transfer enhancement, however lead to a significant pressure drop. The presence of ribs produces typical flow patterns, which disrupts the laminar sub-layer and instigates boundary layer recreation, hence leads to heat transfer enhancement yet with additional pressure drop. In addition to pressure drop penalty, conventional (square/rectangular) rib turbulators possesses more affinity towards the hot-spots at the rear corners of the ribs, as well as at the juncture of ribs and side walls where corner eddies are formed and the flow becomes almost stagnant. These hot-spots can even lead to thermal failure in gas turbine blades. Thus, a large number of rib shapes with and without perforation have been tested which can overcome the problem of hot spot generation (Ali et al., 2016; Buchlin, 1993; Sharma et al., 2017; Tariq et al., 2004). Figure 1 shows the hot spot formation at the rear concave corner of square rib turbulators. Armellini et al. (2008) and Coletti et al. (2008) conducted both experimental and numerical study of a trapezoidal cross-section cooling duct having ribs on one wall and crossing jets. It had been found that the combined action of ribs and jets increased turbulent mixing and hence led to heat transfer enhancement. Shin and Kwak (2008) analyzed the heat transfer performance of ribbed cooling passage of gas turbines blade model with five different kinds of ribs. The results showed that arrangement of holes in staggered manner gave highest enhancement of heat transfer, yet friction factor increased greatly. They suggested to optimize the hole shape for getting maximum thermal performance of the perforated blockage. Ahn et al. (2007) performed an experiment using naphthalene sublimation technique to study mass transfer in rectangular channel with ribs having round and elongated holes. The hole aspect ratios, for each hole to channel area ratio, was analyzed. It was found that, ribs with elongated holes performed better than round

holes in terms of mass transfer. Chung et al. (2013) carried out an experiment determine the effects of improved hole arrays on thermal hydraulic performance of cooling channels. They concluded that an inclined hole array represented improved thermal hydraulic performance over the conventional center hole configuration. Liou et al. (1998) investigated the effect of permeable ribs on the heat transfer performance in the rectangular duct, the results showed that permeable ribs lead to abridged hot spot regions which usually occurred behind the solid ribe, hence leading to significant heat transfer enhancement. Hwang and Liou (1994) reported that the inclination angle of perforation modified the transport nature of flow through jets greatly. The higher inclination angle of hole makes transport properties of the flow more effective. Thianpong et al. (2012) carried out heat transfer and thermal performance study of heat exchanger tube fitted with perforated twisted-tapes. They found that perforated twisted tapes caused a considerable increase in heat transfer rate in comparison to traditional twisted tapes. Jeng et al. (2013) experimentally studied the heat transfer performance of an 180^{o} round turned channel with perforated divider. The authors found that by adjusting the size and angle of perforation, these perforations could achieve spatial thermal regulation and enhance the total heat transfer rate. Qayoum and Panigrahi (2018) carried out an experimental investigation of two-pass channel with permeable ribs using liquid crystal thermography technique. The researchers observed a tremendous enhancement of heat transfer with split slit rib configuration, without any appreciable increase in pressure drop.

As the flow through channels with rib turbulators is highly complex and the internal nature of heat transfer coefficient, friction factor and hot spot formation are primarily controlled by the production and manipulation of large scale vortical structures induced due to modification in rib geometry. Therefore, it is of prime importance to get the detailed spatial and temporal information to understand the heat transfer and fluid flow mechanisms. Recent development in computational fluid dynamics based softwares allows engineers to acquire high resolution temporal and spatial information of temperature and flow field. In this regard Comsol 5.3a can be seen as a promising computational tool to get the detailed information of heat transfer characteristics in the pertinent field.

Pertinent literature as summarized above has scrutinized that perforated blockages can possibly improve hot spots and thermal hydraulic performance of the internal cooling passages. However, the above perforated ribs heat transfer performances has been limited to single pass channel only and no investigations have yet been reported in two-pass channels. Also, in present study square perforation is introduced to study the impact of hole shape on heat transfer and pressure drop characteristics. The primary purpose of this study is to identify the best possible rib perforation within the range of parameter under investigation (i.e. Re) for enhancement of heat transfer at reduced pressure penalty, with

an aim of obviating the possible local hotspots. Based on the heat transfer coefficient and pressure drop measurements, the thermal performance factor for all the rib geometries are assessed in the current computation while utilizing the constant pumping power criteria (Webb and Eckert., 1972).

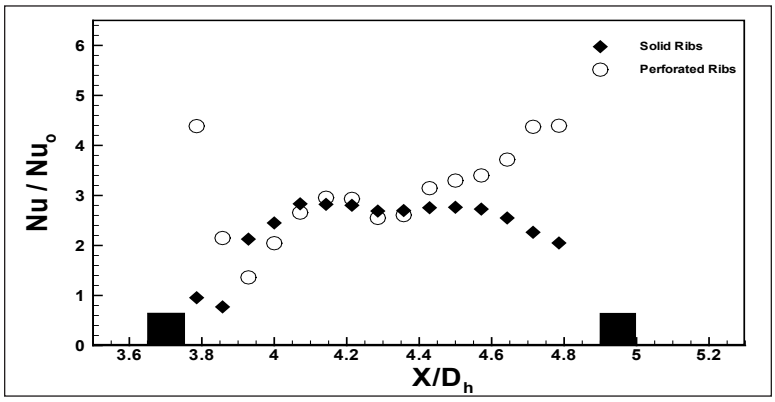


Figure 1. Normalized Nusselt number distribution behind solid and perforated rib turbulators

Numerical Simulation and Physical Model

Figure 2 shows the schematic diagram of computational domain used for present simulation. All the dimensions of the geometry are in mm. The present computational study includes all rib configurations (Case 1-5) as depicted in Table 1. However, only mesh of rib with cylindrical hole is presented in Figure 3. All the four tested perforated ribs differ only in geometry keeping the cross-sectional area same. The length (*l*) of the test section is 1000 mm and the cross-section is 140 mm x 140 mm. The bottom wall is heated with a constant temperature of 100^{o}C for all the summary of cases studied. Ribs were periodically mounted on bottom heated wall, 5 in each pass. The pitch-to-rib height ratio (p/e) and the blockage ratio (e/D$_H$) of ribs are 8 and 0.15, respectively.

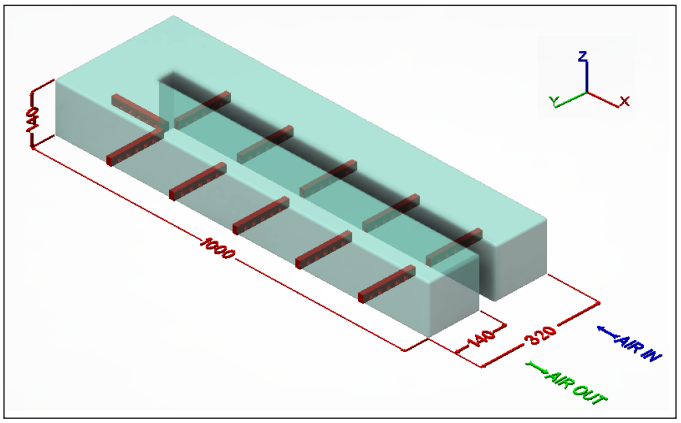
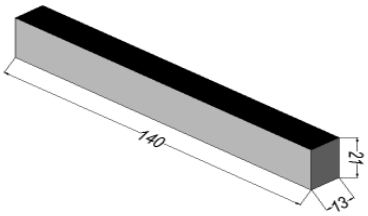
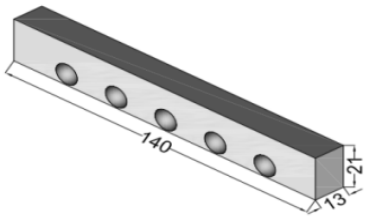
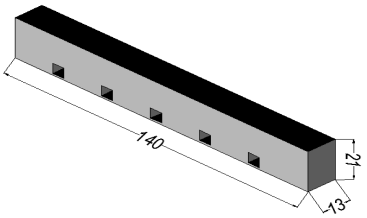
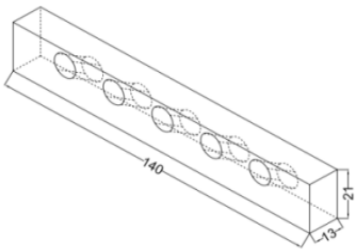
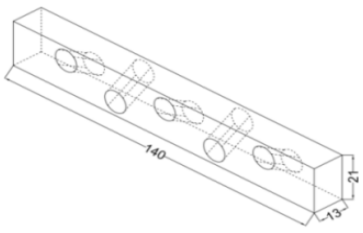
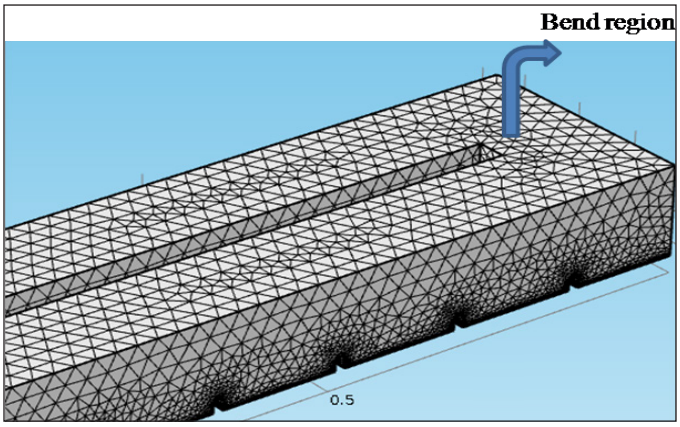


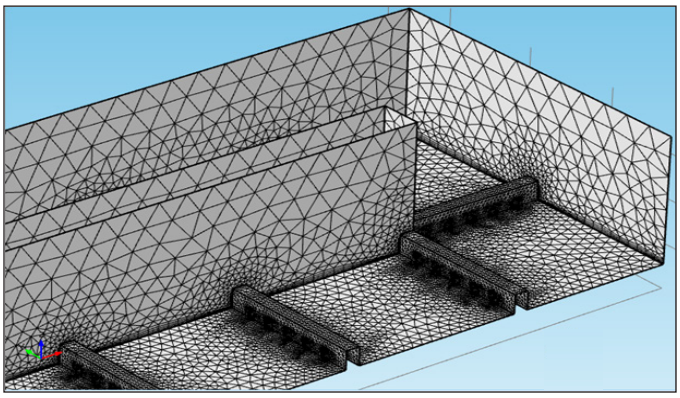
Figure 2. Geometry of a computational domain

Table 1
Construction details of different perforated rib configurations for current study

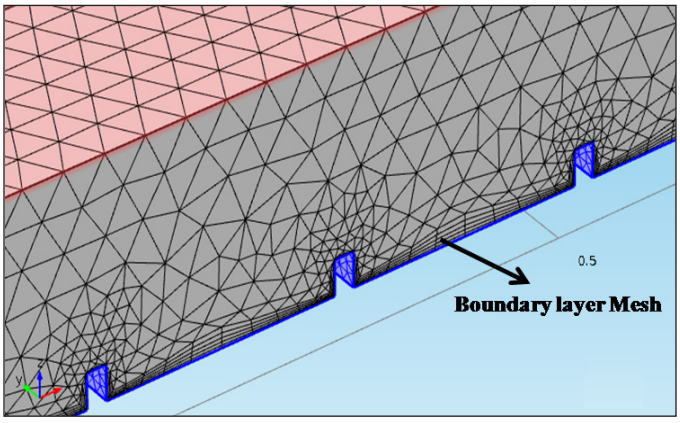
Test Cases	Rib Configuration	Rib Angle
Case 1		Solid Rib
Case 2		Cylindrical hole at bottom
Case 3		Square hole at bottom
Case 4		Cylindrical hole inclined at 30^{o}
Case 5		Tilted cylindrical holes



(a)



(b)



(c)

Figure 3. (a) Overall view of the mesh and (b) Cut-section view of mesh for perforated ribs (c) Enlarged view of boundary layer mesh

Details of Computational Module and Procedure

The present numerical technique uses k-ε module of the commercial software Comsol 5.3a to simulate mean flow characteristics for turbulent flow conditions under specified boundary conditions. The k-ε module is a two-equation model, which by means of two Partial differential equations gives a broad portrayal of turbulence. The k-ε turbulence model provides accurate results of the physics and characteristics of most of the flows in industries, also it is found that it can predict the secondary flow and strong streamline curvature flows correctly. Therefore, the k- ε turbulence model is found to be favorable for the complex fluid flows in the ribbed ducts (Comsol user's guide 5.3a). Comsol 5.3a uses finite element method (FEM) to find approximate solutions of partial differential equations (PDE) and integral equations. A non-isothermal flow has been used where the fluid flow interface is coupled with the heat transfer interface, and pressure forces and the viscous dissipation rate of the fluid are considered. The summary of the steps the above numerical model used are as discretization of the continuous domain and coupled governing equations (elements type and size), generation and propagation of absolute and relative tolerances or errors for convergence criteria, formulating the nonlinear settings for iteration procedure and selection of the suitable solver approach.

In the current study a computational domain was meshed using simple linear free triangular mesh. Figure 3 depicts the overall and enlarged view of the mesh of computational domain. Same size of grid was used for all the cases studied, so as to ensure the homogeneity. The distance of nearest nodes from the wall was maintained to (y+) ~ 11.225 in all the computations. Since the present model has varying geometrical and velocity scales, then consequently the grid and near-wall refinement will have varying levels of resolution. Hence, to have consistency in modelling approach the scalable wall function displaces the near-wall mesh to a y+ value of 11.225 which is the transition to the log-law region. This is necessary since epsilon-based models are not ideal for modelling the laminar sub-layer and the scalable approach internally adjusts the mesh to ensure this region is not resolved.

The current simulation was carried out using parametric segregated approach. The biggest benefit to the parametric segregated solver approach is that it uses the optimal iterative solver in each linear sub-step. The segregated solver normally need additional iterations until convergence, however every iteration takes considerably less time than one iteration of the fully coupled approach. This method involves the integration of the governing equations for mass, momentum, energy and turbulence within the computational domain to build algebraic equations for every dependent variable which is unknown. Earlier investigations delineated that RANS based models could accurately predict the heat transfer and fluid flow characteristics in a two-pass channel.

Governing Equations

In this study, air is taken as incompressible with constant physical properties. Present study uses Reynolds averaged Navier-stokes (RANS) equations and energy equations which can be written in the general form as

Continuity Equation:

$$\rho\nabla.(u)=0 \tag{1}$$

Momentum equation:

$$\rho(u.\nabla)u = \nabla.\left[-pI + (\mu + \mu_T)\nabla u + (\nabla u)^T - \frac{2}{3}\rho kI + F\right] \tag{2}$$

where,

$$\mu_t = \rho C_\mu \frac{k^2}{\varepsilon}$$

Energy equation:

$$\rho C_p u.\nabla T = -\nabla.(k\nabla T) + Q \tag{3}$$

Also, k-ε model adds two additional transport equations and two dependent variables: the turbulent kinetic energy (k), and the dissipation rate of turbulence energy (ε),

Turbulent kinetic energy equation (k):

$$\rho(u.\nabla)k = \nabla.\left[\left(\mu + \frac{\mu_T}{\sigma_k}\right)\nabla k\right] + p_k - \rho\varepsilon \tag{4}$$

Dissipation equation (ε):

$$\rho(u.\nabla)\varepsilon = \nabla.\left[\left(\mu + \frac{\mu_T}{\sigma_\varepsilon}\right)\nabla\varepsilon\right] + Ce_1\frac{\varepsilon}{k}p_k - Ce_2\rho\frac{\varepsilon^2}{k} \tag{5}$$

where,

$$p_k = \mu_T\left[\nabla u : \left(\nabla u + \nabla u^T\right)\right] + \frac{2}{3}\rho k\nabla.u$$

$$C\mu = 0.09 \quad \sigma_k = 1 \quad \sigma_\varepsilon = 1.3 \quad Ce_1 = 1.44 \quad Ce_2 = 1.92$$

Boundary Conditions

The incompressible dry air at ambient conditions is selected as working fluid with constant properties. The fluid flow is steady, three-dimensional, and turbulent. Inlet temperature is considered to be uniform at 300K. Inlet fluid velocity was determined by the Reynolds number. Turbulence intensity (5%) and turbulence length scale (7% of channel hydraulic diameter) were selected for the computation. The boundary condition at inlet was set as velocity inlet and boundary condition (outflow) was imposed at the outlet. A constant temperature boundary condition was used for the ribbed bottom surface with other surfaces being adiabatic and no-slip.

Grid Independence Study

By definition, a grid independent solution means, to make the mesh fine in each spatial dimension and run the simulation again. If the solution for the original and the refined mesh are identical, or nearly so, the solution can be assumed to be grid independent. Grid independence study was carried out with smooth two-pass channel with grid consisting of 200000, 250000 and 330000 elements. The values of averaged normalized Nusselt number found by utilizing above grid sizes are compared in Figure 4. The difference in values of averaged normalized Nusselt number using 330000 and 200000 mesh elements is 6.3%. Thus, to maintain balance amid computational time economy and efficiency, a grid size consisting of 250000 elements was chosen for the present analysis. Few areas are finely meshed for better analysis of the results. These are areas of heated bottom ribbed wall.

Data Reduction

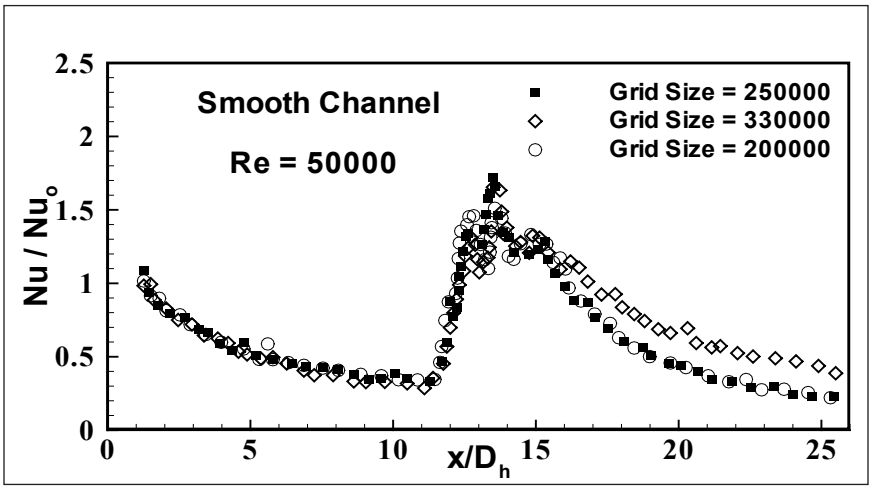


Figure 4. Centerline variation of normalized Nusselt number along Stream-wise direction for three grid sizes

The equation used for determination of heat transfer coefficient over the ribbed surface is as follows:

$$\frac{q''}{(T_w - T_f)} \tag{6}$$

where, q″ stands for heat flux from the heated surface to the surrounding bulk fluid. Velocity of fluid and variations in turbulence are very small and even disappear near the wall. Heat conduction by molecules transports thermal energy in the viscous sub-layer region, which makes the temperature to vary linearly with the distance from the wall (Han and Goldstein., 2008). Hence, the heat flux can be calculated by:

$$q'' = k\frac{\partial T}{\partial n} \tag{7}$$

where, n stands for normal to the surface. Hence, heat transfer coefficient can be calculated from wall temperature, bulk temperature of fluid and heat flux. The local mean bulk fluid temperature is a weighted arithmetic average temperature of fluid and is determined at each stream-wise location using energy balance.

$$Nu = \frac{hD_H}{k} \tag{8}$$

The Nusselt number is normalized by the Dittus- Boelter equation for fully developed turbulent flow in smooth channel.

$$Nu_o = 0.023(\mathrm{Re})^{0.8}\,\mathrm{Pr}^{0.4} \tag{9}$$

Properties of air were calculated at the mean fluid temperature. Friction factor (f) was determined by equation below:

$$f = \frac{\Delta P D_h}{2 l u_{in}^2 \rho_a} \tag{10}$$

Friction factor is normalized by using Blausius equation given below

$$f_o = 0.316/(\mathrm{Re})^{0.25} \tag{11}$$

Performance Index

The performance evaluation of heat transfer enhancement is carried out in order to determine whether it is beneficial to use augmented surface instead of unaugmented case. In present study augmentation cases involve use of surface mounted ribs in the form of solid ribs, and ribs with square and round perforations at different angles. The relation for performance evaluation given by Gee and Webb (1980) is widely employed for comparison of heat transfer performance of a ribbed duct with smooth duct. This thermal performance index is used under constant pumping power criteria. Higher the value of this parameter, higher is the enhancement in heat transfer at small pressure drop. The thermal performance index (η) is expressed as.

$$Nu = \frac{hD_H}{k} \tag{12}$$

Since, a number of fundamental studies has been conducted with various rib geometries, however, the trend of using square ribs is still used for most of the practical applications, hence can be used for the performance evaluation of the new rib geometries. Some factors, i.e., Nu_{ef}, f_r and η_{ef} are defined below to relate the performance indexes of the perforated ribs with the square rib:

$$Nu_{ef} = \left(\left(Nu_{perforated} - Nu_{solid}\right) / Nu_{solid}\right) \times 100 \tag{13}$$

$$f_r = \left(\left(f_{solid} - f_{perforated}\right) / f_{solid}\right) \times 100 \tag{14}$$

$$\eta_{ef} = \left(\left(\eta_{perforated} - \eta_{solid}\right) / \eta_{solid}\right) \times 100 \tag{15}$$

RESULTS

Following section shows the computational results of heat transfer and fluid flow analysis in a stationary two pass channel with blockages (Cases 1-5) at specified Reynolds number range. Detailed normalized Nusselt number and temperature contours are represented for (Cases 1-5) at the range of Reynolds number studied. All the cases are compared for their heat transfer characteristics on the basis of their enhancement of averaged Nusselt number Nu/Nu_o. Characteristics of fluid flow in the duct are explained for understanding the mechanism of heat transfer enhancement by the perforated ribs. At last, comparison is also made with square ribs in terms of Nusselt number enhancement, friction factor and thermal hydraulic performance.

Validation of Model

Before carrying out the simulation for a two-pass channel with ribs, few validation runs have been conducted for smooth channel. Figure 5 depicts the comparison of results for distribution of normalized Nusselt number obtained from present simulation with that of Ekkad et al. (1997), Erille et al. (2015) and Jang et al. (2001). Figure 5 shows the close agreement of the present computational results with that of the experimental results of Erille et al., (2015). However at the beginning of the second pass and hereafter lesser value of heat transfer is found. This variation may be because of the different thickness of divider wall and geometry of the bend section than that of Ekkad et al. (1997), Erille et al. (2015) and Jang et al. (2001). These variations bring about a large variation in the flow conditions which directly affects the heat transfer characteristics. The trend of results from the above computation shows close conformity to the previous experimental and computational results. Therefore computation was carried out with the selected turbulence model.

Overall Heat Transfer and Pressure Drop

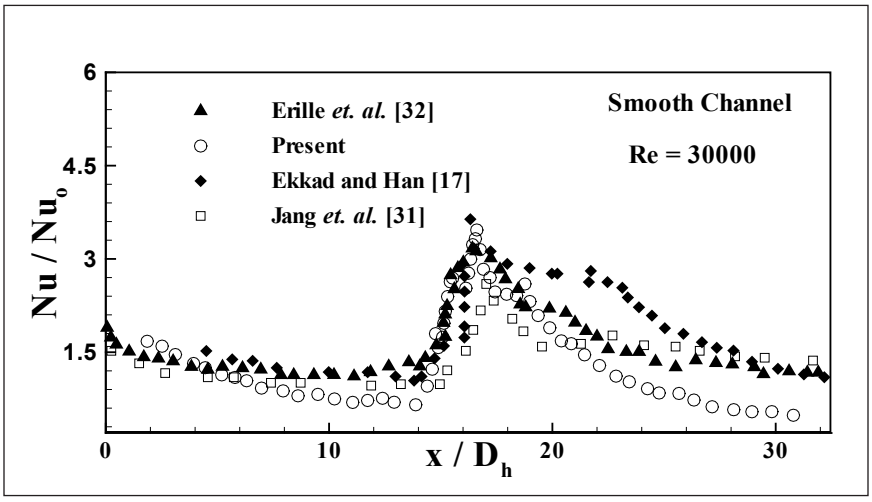


Figure 5. Comparison of normalized Nusselt number with numerical and experimental results

The heat transfer coefficient distribution for perforated rib (Cases 1-5) at Reynolds number range (10000- 50000) and at fixed p/e ratio of 8 have been studied. The detailed analysis of the results reveals that the inter-rib heat transfer distribution is strongly dependent on type of perforation and Reynolds number. To study the fluid flow and heat transfer characteristics of rib roughened channel, it is necessary to understand the secondary flows induced by different perforated rib cases. The fluid flow behavior and turbulence kinetic energy are the major factors that affect the local heat transfer coefficient (Rau et al., 1998). Figure 6 shows the streamline pattern with surface temperature distribution between 3rd and 4th rib of two-pass channel, each perforated case produces a different streamline mapping which

reflects the turbulence features associated with perforated ribs. Typical mapping of the temperature behind solid and perforated rib turbulator is depicted in Figure 7 for Reynolds number equal to 20000. In comparison to solid ribs a significant reduction in temperature at hot spots can be found in case of perforated ribs. Typical temperature contours for (Cases 1-5) is shown in Figure 8.

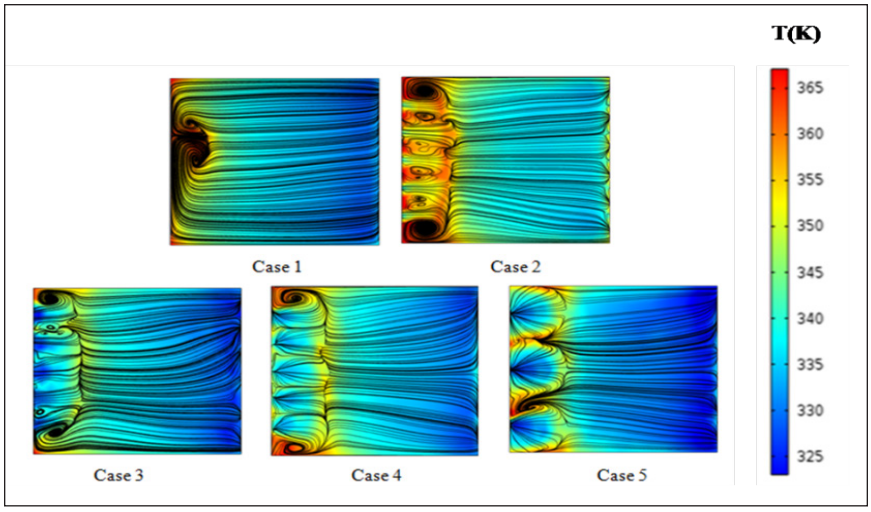


Figure 6. Streamlines with surface temperature distribution for perforated ribs

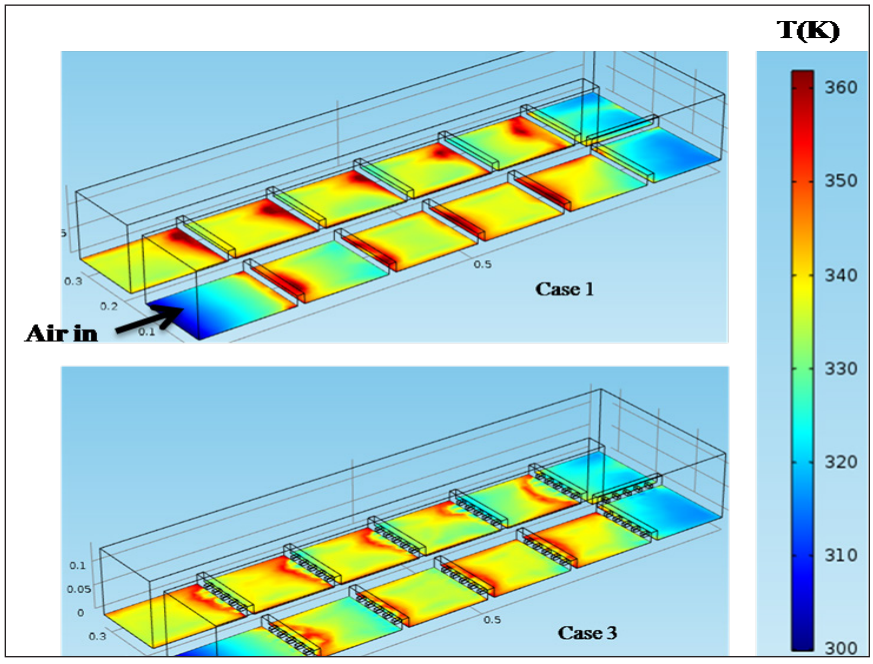


Figure 7. Temperature contours for conventional solid and perforated ribs.

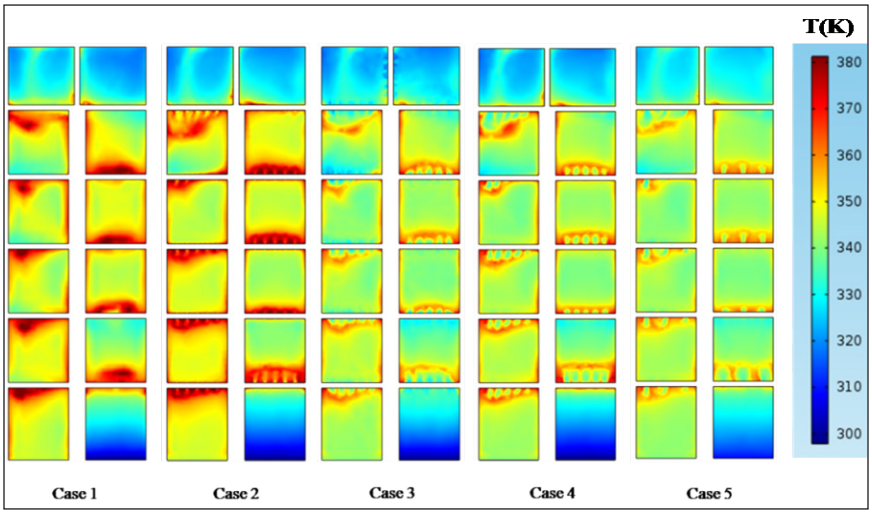


Figure 8. Temperature contour of solid and perforated ribs

Figure 9 provides the mappings of Nu/Nu_o for (Cases 1-5) for the lowest Reynolds number. The contour plots clearly show the distinct distribution of Nu/Nu_o corresponding to different perforation cases at same Reynolds number and p/e ratio. Present study displays the effect of perforations towards obviating local hot spot regions, which typically exists just behind the square rib. For most of the Reynolds number perforated ribs provide higher value of normalized Nusselt number than square rib. The modification in heat transfer enhancement can be observed in all the perforated rib cases. This variation in heat transfer enhancement can be possibly because of air flowing through the perforation, causing a modification in secondary flow which breaks the vortices between the ribs. Case 2 ribs gives rise to the formation of a multi wall-jet system, the strength of which decreases quickly so that no more effect is appreciated at a downstream distance. From this location no significant variation of Nusselt number is found indicating that the flow has retrieved its two-dimensional character. However by increasing the perforation angle to 30^o fresh air reaches the base of heated plate with more intensity leading to enhancement in heat transfer considerably. Case 3 and Case 5 each have different cooling results. Case 3 ribs show highest heat transfer enhancement, as fluid jets are comparatively very near to the side walls, followed by the Case 5 and Case 4. This becomes clearer when looking at Figure 10 which depicts the heat transfer distribution for (Cases 1- 5).

Figure 11 shows the span-wise averaged normalised Nusselt number ratio at different Reynolds number. The area averaged Nusselt numbers are determined by averaging the local Nusselt numbers over the heated surface on the bottom wall of the channel. Case 2 shows lowest value of Nusselt number augmentation, because the strength of multi wall-jet system formed by the hole at no inclination, decreases quickly so that no more effect is appreciated

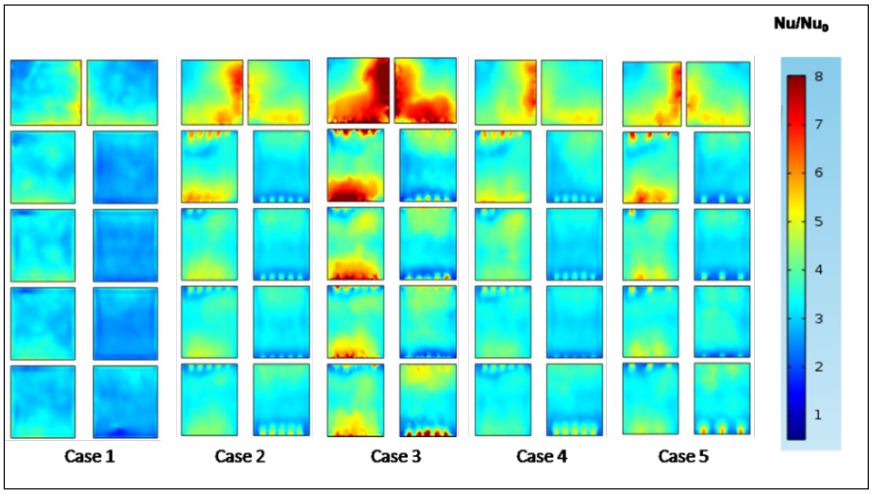


Figure 9. The normalised Nusselt number contours of solid and perforated ribs

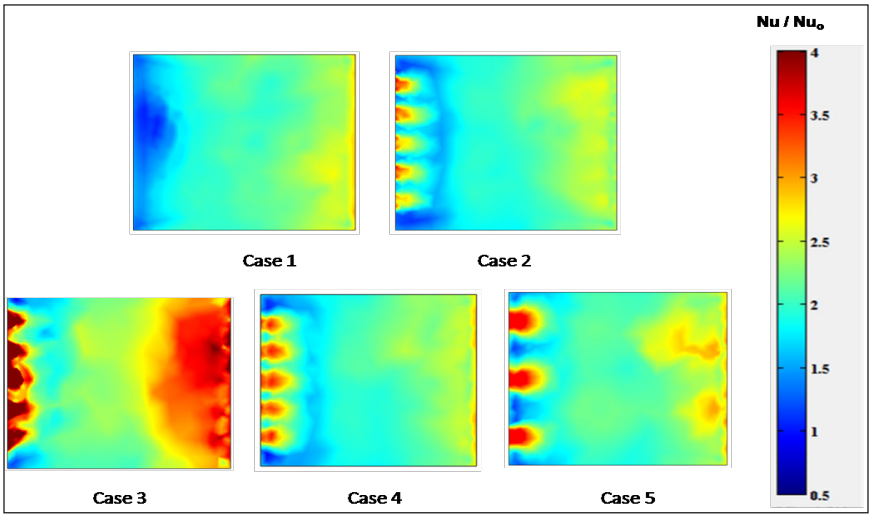


Figure 10. Normalised Nusselt number distribution between third and fourth rib behind perforated ribs

at a downstream distance. The fact that more inclination brings more enhancements is discussed in Figure 11. As depicted in the figure the performance of perforations increases from 0^o to 30^o and tilted perforations further enhance the normalized Nusselt number, the possible reason could be due to the pointing of jets from both sides of the ribs to the bottom heated wall of the channel. Further analysis of Figure 11 depicts that Nu/Nu$_o$ is a powerful function of Reynolds number. It decreases with the increase in Re from the maximum at 10000 to the lowest at 50000, which is a well-known behaviour of ribbed ducts (Ekkad et al., 1997; Erelli et al., 2015).

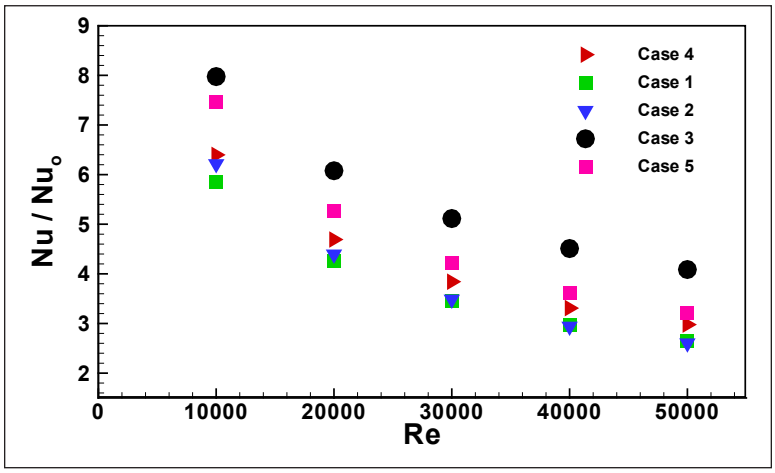


Figure 11. Normalised Nusselt Number Nu/Nu_o at various Re

The ribs installed in two-pass channel acts as an obstruction to the flow and lead to separation of flow, recirculation and reattachment, therefore leading to the higher pressure loss, and hence higher friction factor for ribbed duct compared to smooth channel at same Reynolds number. The perforated ribs shows higher pressure drop as compared to square ribs except for Case 2 at the present Reynolds number range Figure 12 shows the computational results for f/f_o as a function of different perforated ribs for different Reynolds numbers at p/e of 8. Friction factor results evidently signify that the installation of ribs lead to pressure drop increment up to 0.89−6.3 % higher than that of the solid ribs under similar flow conditions. Clearly, friction factor is the performance parameter related to pressure penalty. The friction factor ratio increases slightly with the increase in Reynolds number. Inclined holes shows highest friction factors than the ribs with other perforations.

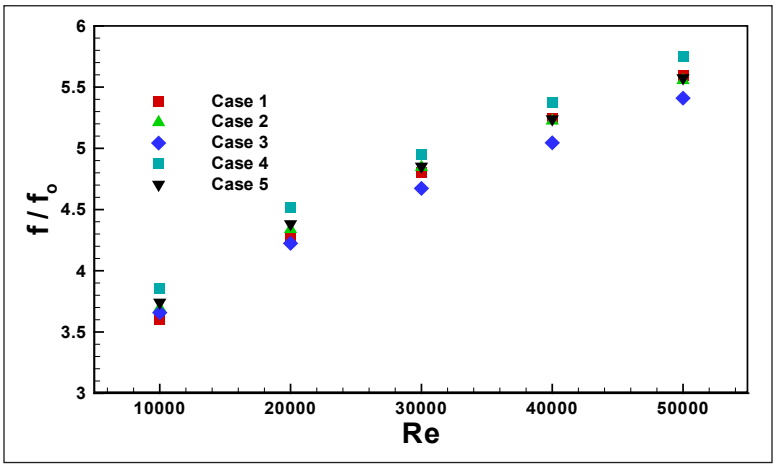


Figure 12. Normalised friction factor f/f_o at various Re

Figure 13 presents the thermal performance factor for (Cases 1-5) calculated using Eq. (12). It is clear from the figure, that thermal performance factor decreased as the Reynolds number increased. Among the tested perforated blockages, Case 3 showed the highest thermal performance factor. Case 1 and Case 2 shows almost equivalent thermal hydraulic performance. Other cases show performance greater than that of the conventional square rib. The difference in thermal hydraulic performance is more prominent at lowest Reynolds number.

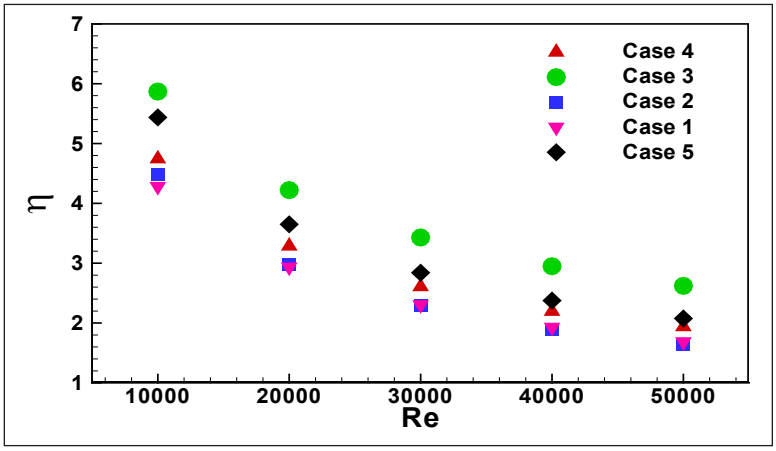


Figure 13. Thermal-hydraulic performances at various Reynolds number

Table 2 shows the comparison of Nu_{ef}, f_r and η_{ef} for different Reynolds number at fixed p/e of 8. For most of the Reynolds number, perforated ribs show better values Nu_{ef}, f_r and η_{ef} than square ribs. On the contrary, at almost all values of Reynolds number except for Case 3, square ribs perform better than perforated ribs from relative friction factor point of view. Also, Case 3 shows highest value of most of the Nu_{ef}, f_r and η_{ef} for all Reynolds number considered, except the performance in terms of friction factor at lowest Reynolds number is found to be less than the square rib.

Table 2
Comparison of relative performance indexes against square rib at different Re for different perforated cases

P/e	Cases	Re				
		10000	20000	30000	40000	50000
		Nu_{EF} f_R η_{EF}	Nu_{EF} f_R η_{EF}	Nu_{EF} f_R η_{EF}	Nu_{EF} f_R η_{EF}	Nu_{EF} f_R η_{EF}
8	Case 2	20.07 2.30 0.31	19.91 1.52 4.07	21.2 0.89 5.86	22.80 -0.36 3.03	23.54 -0.69 0.99
	Case 3	28.3 1.35 20.92	38.30 -0.93 52.77	46.7 -2.40 53.05	53.40 -3.50 49.52	57.05 -3.09 55.10
	Case 4	-5.80 6.30 5.33	-3.28 5.40 14.55	-0.94 2.90 27.51	0.73 2.10 29.93	1.40 2.40 32.39
	Case 5	2.94 0.40 3.88	6.72 0.40 21.34	10.23 1.02 32.50	12.60 -0.12 33.68	14.40 -0.38 34.99

CONCLUSION

Numerical simulations for steady-state heat transfer and pressure drop measurements behind two dimensional perforated ribs in a two-pass channel were performed. Average Nusselt number and overall friction factor results were obtained for Cases 1 to 5 at different Reynolds numbers. The area of perforations was kept same. The mapping of the heat transfer coefficient provides a useful visualization support to interpret the complexity of the resulting flow. The results are summarized below for the conditions and geometries under which the simulations were conducted:

1. The study compared the average heat transfer coefficients obtained with five perforation designs in rib turbulators. Fraction of the air swept through the perforated rib and interacted with the recirculation bubble generated behind the rib by the impingement of the upper free stream on the wall. As a consequence, the hot spot area occurring in the region just behind a solid-type rib did not exist anymore in the corresponding region of the perforated turbulator.
2. Perforations in the blocks enhanced the heat transfer coefficient on the wall segments between the ribs by 1.4-57 % than that of the flow through the channel with conventional solid ribs at the same flow rate.
3. Inclined perforations in ribs performed better in terms of heat transfer coefficients. Depending upon the inclination angle, inclination may lead to 18.3 to 25.4% increase in heat transfer performance than the conventional solid ribs.

4. Perforated blocks led to larger pressure drop as compared to conventional square blocks.
5. For a given pumping power, Case 3 generated highest enhancement of heat transfer than that of the square ribs.

REFERENCES

Acharya, S., Dutta, S., Myrum, T. A., & Baker, R. S. (1993). Periodically Developed Flow and Heat Transfer in a Ribbed Duct. *International Journal of Heat and Mass Transfer, 36*(8), 2069–2082.

Ahn, H. S., Lee, S. W., Lau, S. C., & Banerjee, D. (2007). Mass (heat) transfer downstream of blockages with round and elongated holes in a rectangular channel. *Journal of Heat Transfer, 129*(12), 1676–1685.

Ali, M. S., Tariq, A., & Gandhi, B. K. (2016). Role of chamfering angles and flow through slit on heat transfer augmentation behind a surface-mounted rib. *J. Heat Transfer 138*(11), 1-16.

Armellini, A., Coletti, F., Arts, T., & Scholtes, C. (2008). Aero-thermal investigation of a rib roughened trailing edge channel with crossing-jets. Part I: flow field analysis. In *Proceedings of ASME Turbo Expo* (pp. 615-661). Berlin, Germany.

Buchlin, J. M. (1993). Heat transfer behind a rectangular cylinder in a boundary layer. ***J****ournal of Applied Fluid Mechanics, 4*(1), 137-149.

Chamoli, S., & Thakur, N. S. (2015). Thermal Behaviour in Rectangular Channel Duct Fitted with V- Shaped Perforated Baffles. *Heat Transfer Engineering, 36*(5), 471- 479.

Chung, H., Park, J. S., Sohn, H. S., Rhee, D. H., & Cho, H. H. (2013). Trailing edge cooling of a gas turbine blade with perforated blockages with inclined holes. In *Proceedings of ASME Turbo Expo 2013: Turbine Technical Conference and Exposition* (pp. V03AT12A043-V03AT12A043). San Antonio, Texas, USA.

Coletti, F., Armellini, A., Arts, T., & Scholtes, C. (2008). Aero-thermal investigation of a rib roughened trailing edge channel with crossing-jets, Part II: heat transfer analysis. In *Proceedings of ASME Turbo Expo* (pp. 663-673). Berlin, Germany.

Ekkad, S. V., & Han, J. C. (1997). Detailed heat and mass transfer distribution in a two-pass Square channel with rib turbulators. *International Journal of Heat and Mass Transfer, 40*(11), 185-193.

Erelli, R., Saha, K., & Panigrahi, P. K. (2015). Influence of turn geometry on turbulent fluid flow and heat transfer in a stationary two-pass square duct. *International Journal of Heat and Mass Transfer, 89*, 667-684.

Gao, B., Bi, Q., & Gui, M. (2016). Experimental Performance Comparison of Shell Side Heat Transfer for Shell-and-tube Heat Exchangers with Different Helical Baffles. *Heat Transfer Engineering, 37*(18), 1566-1578.

Gee, D. L., & Webb, R. L., (1980). Forced convection heat transfer in helically rib roughened tubes. *International Journal of Heat and Mass Transfer, 23*(8), 1127–1135.

Han, J. C., Dutta, S., & Ekkad, S. (2012). *Gas turbine heat transfer and cooling technology*. New York, USA: Taylor and Francis Group, CRC Press.

Han, J. C., Park, J. S., & Lei, C, K. (1985). Heat Transfer Enhancement in Channels with Turbulence Promoters. *Journal of Engineering for Gas Turbines and Power, 107*(3), 628-635.

Han, S., & Goldstein, R. J. (2008). The heat/mass transfer analogy for a simulated turbine blade. *International Journal of Heat and Mass Transfer, 51*(21-22), 5209-5225.

Jang, Y. J., Chen, H. C., & Han, J. C. (2001). Computation of flow and heat transfer in two pass channels with 60^{o} ribs. *Journal of Heat Transfer, 123*(3), 563–575.

Jeng, T. M., Tzeng, S. C., & Xu, R., (2013). Fluid visualization and heat transfer tests in an 180^{o} round turned channel with a perforated divider. *International Communications in Heat and Mass Transfer, 44*, 45-51.

Liou, T. M., & Chen, S. H. (1998). Turbulent Heat and Fluid Flow in a Passage Disturbed by Detached Perforated Ribs of different heights. *Journal of Heat Mass Transfer, 41*(12), 1795-1806.

Liou, T. M., & Hwang, J. J. (1994). Augmented Heat Transfer in a rectangular channel with permeable ribs mounted on one wall. *Journal of Heat Transfer, 116*(4), 912-920.

Liou, T. M., Chen, C. C., & Tsai, T. W. (2000). Heat Transfer and Fluid Flow in a square duct with 12 different shaped vortex generators. *Journal of Heat Transfer, 122*(2), 327-335.

Multiphysics, C. O. M. S. O. L. (2014). *Heat transfer module user's guide. COMSOL version, 4*. Burlington, MA: COMSOL Inc.

Qayoum, A., & Panigrahi, P., (2018). Experimental investigation of heat transfer enhancement in a two-pass square duct by permeable ribs. *Heat Transfer Engineering,* 1-2. doi: 10.1080/01457632.2018.1436649

Rau, G., Cakan, M., & Moeller, D. (1998). The Effect of Periodic Ribs on the Local Aerodynamic and Heat Transfer Performance of a Straight Cooling Channel. *ASME, 120*(2), 368–375.

Sharma, N., Tariq, A., & Mishra, M., (2017). Detailed Heat Transfer Investigation inside a Rectangular Duct with an Array of Ventilated Rib Turbulators. In A. K. Saha, D. Das, R. Srivastava, P. K. Tanigrahi & K. Muralidhar (Eds.), *Fluid Mechanics and Fluid Power –Contemporary Research* (pp. 775-784). India: Springer.

Shin, S., & Kwak, J. S. (2008). Effect of hole shape on the heat transfer in a rectangular duct with perforated blockage walls. *Journal of Mechanical Science Technology, 22*(10), 1945– 1951.

Tariq, A., Panigrahi, P. K., & Muralidhar, K. (2004). Flow and Heat Transfer in the Wake of a Surface Mounted Rib with a Slit. *Experiments in Fluids, 37*(5), 701–719.

Tariq, A., Singh, K., & Panigrahi, P. K. (2003). Flow and Heat Transfer in Rectangular Duct with Single Rib and Two Ribs Mounted on the Bottom surface. *Journal of Enhanced Heat Transfer, 10*(2), 171-198.

Thianpong, C., Eiamsa-ard, P., & Eiamsa-ard, S., (2012). Heat transfer and thermal performance characteristics of heat exchanger tube fitted with perforated twisted-tapes. *Heat Mass Transfer, 48*(6), 881-892.

Webb, R. L., & Eckert, E. R. G. (1972). Application of Rough Surfaces to Heat Exchanger Design. *International Journal of Heat Mass Transfer,15*(9), 1647– 1658.

Webb, R. L., (1994). *Principles of enhanced heat transfer*. New York, USA: John Wiley and Sons.

Yuan, Z. X., Tao, W. Q., & Yan, X. T. (2003). Experimental Study on Heat Transfer in Ducts with Winglet Disturbances. *Heat Transfer Engineering, 24*(2), 76-84.

APPENDIX

Nomenclature

D_h	Hydraulic diameter	[m]
e	Rib height	[m]
F	Body force vector	[N/m3]
f	Friction factor	
f_r	Friction factor reduction compared to square rib	
h	Heat transfer coefficient	[W/m^{2}K]
K	Thermal conductivity	[W/mK]
L	Channel length	[m]
Nu	Nusselt number	
Nu_{ef}	overall Nusselt number enhancement factor compared to square rib	
P	Rib pitch	[m]
p	Pressure	[Pa]
q″	Heat flux	[W/m^2]
Re	Reynolds number	[$\rho\ u_{in}\ D_h/\mu$]
u	Fluid velocity	[m/s]

Subscripts

f	Fluid
in	inlet
o	From Dittus-Boelter Coorelation
w	wall

Greek symbols

ρ_a	Density of air [kg/m^3]
η	Thermo-hydraulic performance
η_{ef}	Factor for performance enhancement compared to square rib
μ	Dynamic viscosity [kg/ms]

Pertanika J. Sci. & Technol. 26 (4): 2031 - 2044 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Design and Development of Processing Apparatus for Young Coconuts

Muhamad Hellmy Hussin

Fabrication & Joining Section, Universiti Kuala Lumpur Malaysia France Institute, Section 14, Jalan Teras Jernang, 43650 Bandar Baru Bangi, Selangor, Malaysia

ABSTRACT

As we know, drinking young coconut water and eating the tender meat give many benefit to the body for its nutritious value rather than its taste; but do we realize that it requires a dangerous tasks in processing it. The process of trimming requires skills which only can be obtained by those who run the work daily. Thus, a portable apparatus which has the capabilities of reducing the hazardous tasks and fasten the time consumed for processing the young coconut fruit has been proposed and developed. The development of the product begins with collecting and analyzing the data of 30 young coconut fruits. Then, it is followed by designing the whole product at main and component level. The conceptual design is done initially using freehand sketching technique. Next, the 3D solid modeling relies totally on the CATIA V5R19 software. Finally, a complete details drawing is produced using CAD software. In this work, the design focuses on the blade slicing and punch bit head to reduce the hazardous tasks during processing of the young coconut. The blade is designed to allow the slicing movement to be maneuvered during the husk removal process. Meanwhile, the puncher has replaced the usage of chopper in creating an opening at the top of the endocarp. Thus, the device developed will reduce the hazardous task by eliminating the chopping process and replacing it with the slicing process. Therefore, the tendency to get caught in accidental injury during the chopping process can be significantly reduced.

Keywords: CATIA, computer aided design, conceptual design, innovation, product design

ARTICLE INFO

Article history:
Received:
Accepted:
Published: October 2018

E-mail addresses:
hellmy@unikl.edu.my

INTRODUCTION

Coconut fruit is among the 20 important crops in the world (Vidhan Sigh & Udhayakumar, 2013). The coconut provides a nutritious source of juice, milk, and oil that has fed and nourished populations around

ISSN: 0128-7680
e-ISSN: 2231-8526

the world for generations (Lihua, 2015). On many islands coconut is a staple in diet and provides the majority of the food eaten (Mani & Jothilingam, 2014). Nearly, one third of the world's population depends on coconut to some degree for their food and their economy. The coconut is made of several layers (Adzimah & Turkson, 2015; Ketan, Vinod, & Sakhale, 2014). The outer layer where it is called the husk (mesocarp) is fibrous and the second layer is an inner "stone" (endocarp) and the third layer is the tender white meat, sort of jelly. The inner part is the cavity which filler with "coconut water" (Venkataramanan, Abhinav Ram, & Rahul, 2014). A very young coconut has very little meat (flesh), and the meat is very tender, almost a gel (Yahya & Zainal, 2014). For thousands of years, coconut products have held a respected and valuable place in local folk medicine (Syafriani, Apriantono, & Sigit, 2014). Meanwhile, modern medical science confirming the use of coconut in treating many conditions such as colds, constipation, cough, fever and flu (Yuniwarti, Asmara, Artama, & Tabbu, 2015).

As well known, drinking young coconut water and eating the tender meat give many benefits to the body for it dangerous tasks in processing it. To consume their flesh meat and the juice, trimmed coconuts must be opened either manually or mechanically (Jarimopas, Ruttanadat, & Terdwongwarakul, 2009). Young coconuts sold in the restaurants, stalls or even in the fresh markets are usually trimmed in a sort of conical shape (Mohanraj, Rahgul Krishna, Kannan, Rajkumar, & Elango, 2014). The process of trimming requires skills which only can be obtained it is done daily and it would be possible only for those who runs the young coconuts business and also for the agricultural biomass industries (Salleh, Mohd Yusoh, & Ruznan, 2015; Sim, Mohd Irwan Lu, Lee, Mohamed, 2015). Thus, for the commoners, it is still a dangerous task (Ismail et al., 2015). Traditionally, in current practice the process of removing the top and bottom husk is by chopping them off. The process of removing the husk really needs a cautious move and focused mind in order to maneuver the sharp and heavy object to cut the husk (Abraham, Girish, Vitala, & Praveen, 2014). Furthermore, the oval shape of the coconut produces instability during the cutting process, thus it adds up the percentage of accidental injuries.

Even though there are already devices which almost have similar functions to the device which is going to be developed but there is still room for improvement (Vinod, Ketan, & Chandrashekhar, 2014). Some of the available products developed are not meant for the small scale entrepreneur or home users (Prashant, Gopinath, & Vignesh, 2014). The existing product in the market is big (for mass production), custom made, not portable, complicated to be operate, expensive and dependent to other sources such as hydraulic, pneumatic and power supply (Sabale & Kolhe, 2015; Sangameshwara & Ravaikiran, 2015; Satip & Kiattisak, 2008). Thus, it is an advantage to design the young coconut processing apparatus that is independent from other sources, portable with optimum size, user friendly, and not too costly compared to the existing product.

It is observed that, the possibility of a non professional chopper tends to get injured during processing the young coconut is extremely high. This dangerous task could invite accidental injuries if it is performed by the non-experts in the field (Vasconcelos & Junior, 2015). Thus, the development of an apparatus could help people to avoid the accidental injuries during processing of the young coconut fruits. Therefore, this paper focuses on designing and developing a portable multi-tasking device which can eliminate and reduce the hazardous task and fasten the time taken on processing the young coconut fruits when compared to the conventional processing steps.

MATERIALS AND METHODS

Design Requirement

The design requirement is derived from the needs of the client such as mechanisms, performance of the client such as the mechanisms, performance parameters, reliability, and safety. These elements will derive and control the design and constraints throughout the process. This is to determine the critical process which can cause the accidental injuries and concurrently will provide ideas during designing stage at product level.

Conceptual Design

The design concept of total physical shape of the young coconut processing apparatus was performed. The design concepts addressed the problem highlighted. In this process, the entire elements such as design requirements, the customer needs and constraints were considered in order to satisfy the needs.

Preliminary Design

In this stage, the overall system configuration is defined, and drawing definition will be developed to provide early project configuration and to assist coordination during the detail design phase.

Data Collection

Before developing the detail design phase, a few crucial data were required to obtain which are, (i) dimension of the young coconut fruits, and (i) the force required in order to punch a hole on the top of the endocarp and to cut the young coconut into halve. The defined dimensions such as, (i) the height of the young coconut fruit before the removal of the top and bottom husk, H1, (ii) the height of the young coconut fruit after the removal of the top and the bottom husk, H2, and (iii) the diameter of the young coconut fruit, D. The data collection was to provide dimensions and measurement that would allow further construction of the detail design on the product component level. The data of 30 samples

of different types, sizes and origin of the young coconut fruits were collected. For the study purposes, the selection of the coconut fruits relied totally on the experienced of the young coconut water sellers. The issue of maturity of the young coconut fruit was not really important as long as it falls under the category of young coconut fruit.

Detail Design

A detail design of each parts of the device at main and sub level were produced. All the manufacturing drawing such as details drawing, 3D drawings, general assembly drawing, general sectional assembly drawing and exploded drawing were constructed. For the details drawing on components and assembly level, 3rd angle orthographic projection were used. This was performed using the parametric modeling software, CATIA V5R19.

Product Modeling

The 3D modeling software CATIA V5R19 was used to model the young coconut processing apparatus. All the freehand sketches of the conceptual and the preliminary design were detailed using this software. The dimensions and form of each component was assessed and defined based on the engineering design analysis and also with Heuristic design technique. The product and its sub element would be simulated concurrently during modeling phase. This process was to determine the feasibility of assemblage and the tolerance needed.

RESULTS AND DISCUSSION

Conceptual Design

A few design concepts were performed. The design concepts were in line with the objectives points of the apparatus which include, (i) the hazardous task of chopping the top and the bottom husk, (ii) cutting the young coconut into halves after the drainage of its water, (iii) creating an access at the top of the endocarp for the young coconut water draining purpose, (iv) the apparatus has still to be operated in manual mode but without the chopping process, and (v) the husk removal process has to assemble like the sawing or slicing process.

Figure 1 shows the preliminary design 5 into sub-units. The main function of the device is, (i) to remove the top husk, (ii) to remove the bottom husk, (iii) to create a hole through the top endocarp, and (iv) to cut the young coconut into two parts. As indicated in Figure 2, the sub-unit 1 shows the Positioner which the aim of this sub-unit is to position the young coconut vertically and horizontally and to provide stability to the young coconut – additional contact point. It is also designed to allow rotation to the young coconut. Sub-unit 2 shows the Base which the objective of this part is to provide rigidity and stability of the device. Meanwhile, sub-unit 3 is the Slicer. It is designed for slicing the husk (top and bottom) and to cut the young coconut into two parts. The sub-unit 4 is the Slicer Pivoting Mechanism

which is designed to provide rotation and sliding angle to the knife and the sub-unit 5 is the Puncher. The role of the Puncher is to create access through endocarp of the coconut.

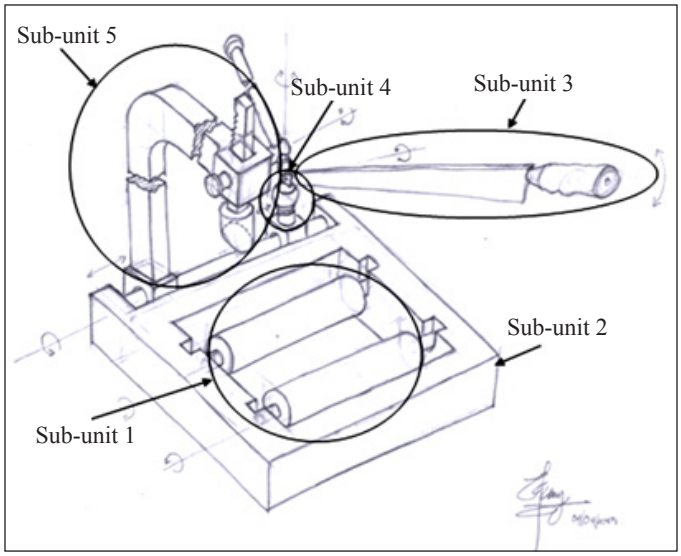


Figure 1. Preliminary design for five (5) sub-units

Data Collection

The data provides the maximum and minimum dimension of H1, H2, and D from the 30 samples of the young coconut fruits. Figure 2 shows the defined dimensions of the young coconut fruit. By rearranging the data, it defines the working area (WA) of the apparatus to be developed as viewed in 3-dimensional views. The WA was presented in Figure 3. The WA will be the constraints of effective working area to the development of the total size of the apparatus. Figure 4 shows the young coconut processing apparatus working area terms definition.

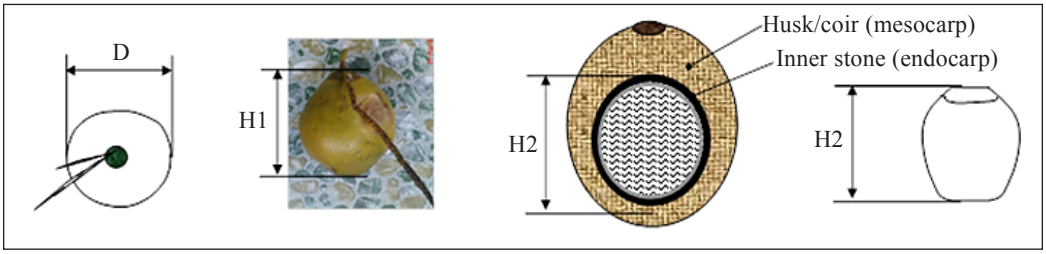


Figure 2. Young coconut fruit defined dimensions

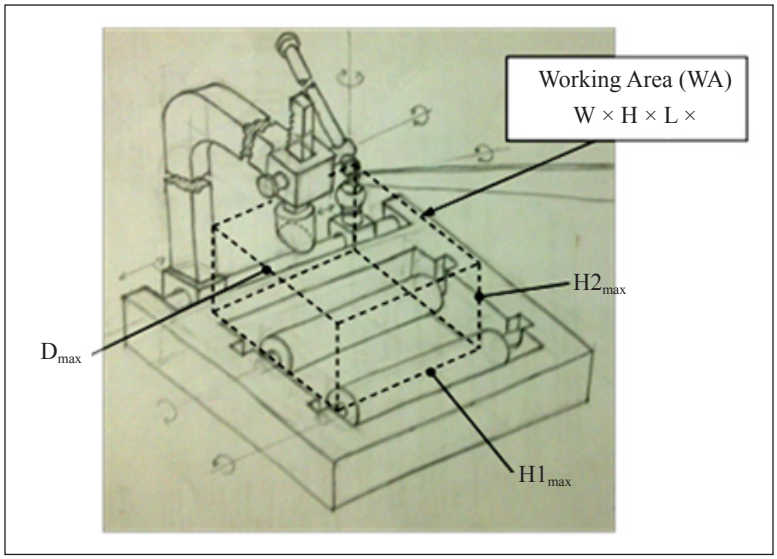
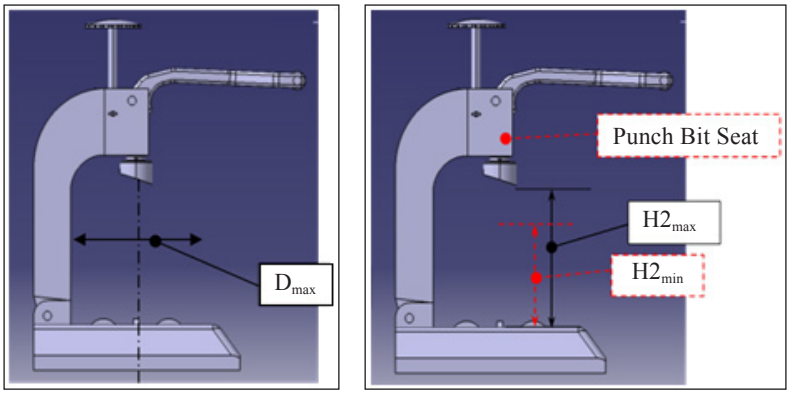


Figure 3. Young coconut processing apparatus's working area



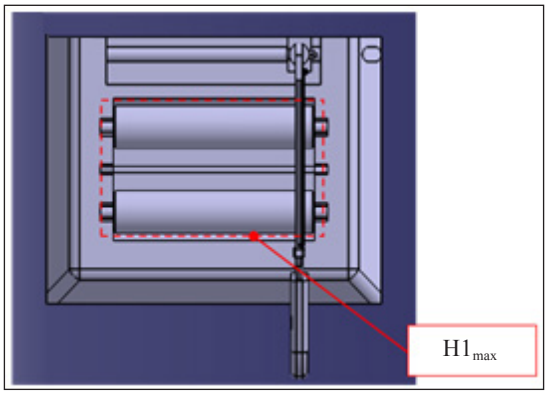


Figure 4. Young coconut apparatus working area terms definition

In Figure 4, the width (W) can be obtained from the dimension of D_{max}; the height (H) can be obtained from the dimension of $H2_{max}$. The length of the WA is obtained from the dimension of $H1_{max}$. The purpose of $H2_{min}$ is to provide the minimum height of the young coconut fruit. This data enabled to design the punch bit head to be lowered to the minimum height (H2) of the young coconut. The difference between $H2_{max}$ and $H2_{min}$ acts as the tolerance range for the punch bit seat to keep functioning.

Punch Bit

The Punch bit part is designed with the slant angle of 18° as viewed in Figure 5. The purpose of the designed punch bit is to reduce the surface area during the penetration to the endocarp of the young coconut which would reduce the pressure on the target surface. In this part, the target surface is located between the tip of the punch bit and the top endocarp of the young coconut. The reduced value of the pressure acts directly on the target surface will reduce the force to be applied during punching process. As indicated in Figure 5, the selected angle of 18 ° provides a small value of the penetration length. The smaller value provides faster complete penetration at the top endocarp of the young coconut fruit.

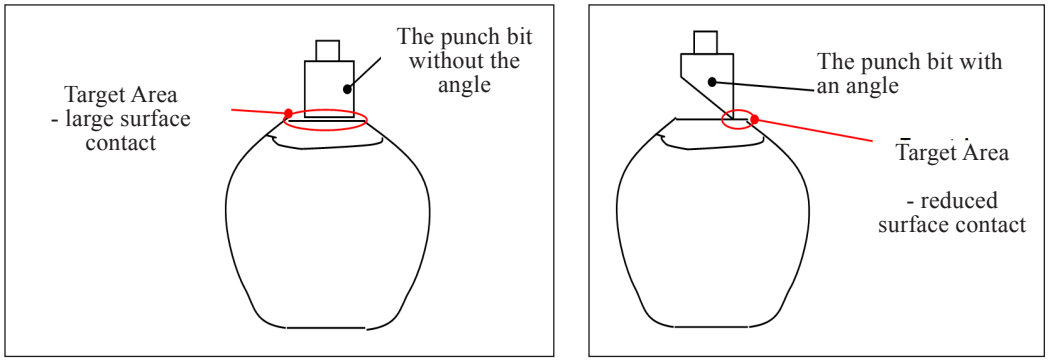


Figure 5. The effective area of a punch bit design

Punch Bit Seat

The punch bit seat is designed in moveable vertical direction. It can be adjusted to accommodate tolerance zone between $H2_{max}$ and $H2_{min}$. As for the punch bit seat not to be disoriented while being lowered, a guide column is embedded in the component. Figure 6 shows the different position of the punch bit seat.

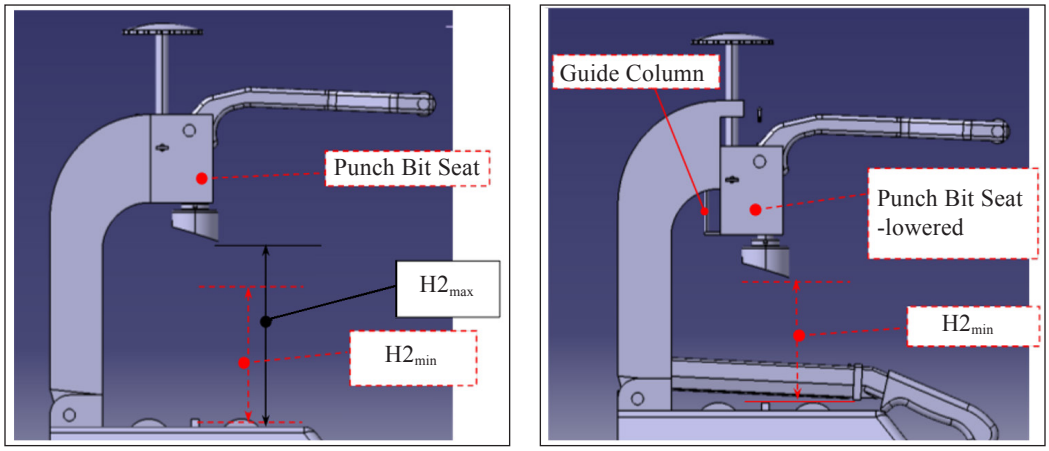


Figure 6. The different position of the punch bit seat

Punch Lever

The designed punch lever in Figure 7, can be used for left and right handed person. When one pressing the punch lever downwards, the roller will rotate with respect to the axis center. This would act as a friction reducer between the palm skin and the roller.

Blade Shutter

As for the safety purpose when handling the sharp blade during cutting and slicing process or when the apparatus is not being used, the sharp blade has to be remained in safety mode. This is to avoid any accidental injuries that can be caused by the sharp blade. Thus, the blade shutter is designed to cover the blade sharpness. The blade shutter is forced upward during the slicing and cutting process while it returns to its initial position when pulled by the gravity which acts on its own mass as shown in Figure 8.

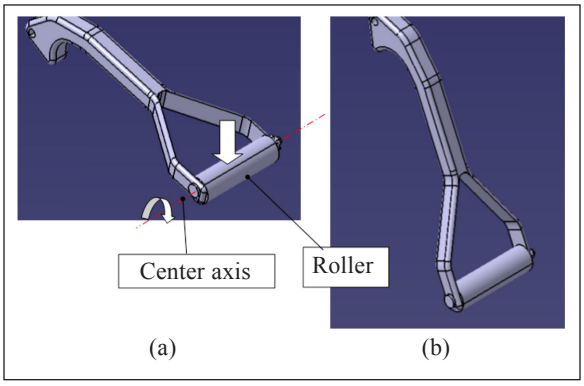


Figure 7. Puncher lever

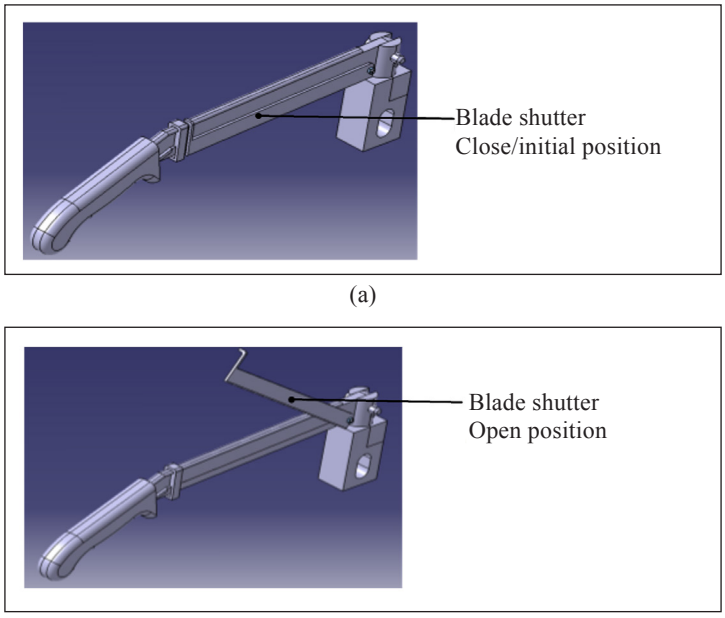


Figure 8. Blade safety mechanism

The angle of the blade holder was designed to comfort during handling the blade. The selection of the angle value is based on the observation of the normal position and condition during holding and grasping an object in standing position. As a result, an angle of 20^{o} was selected to mimic the position of hand grasping or holding and object pointing outward as shown in Figure 9.

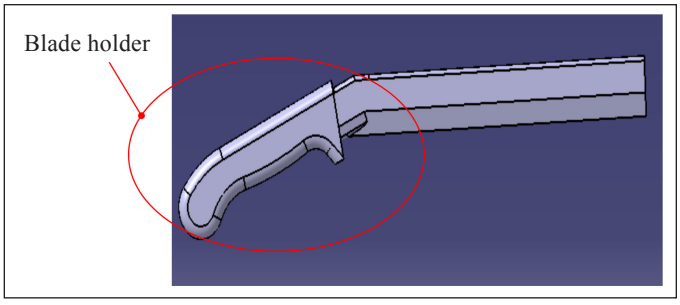


Figure 9. Blade holder design

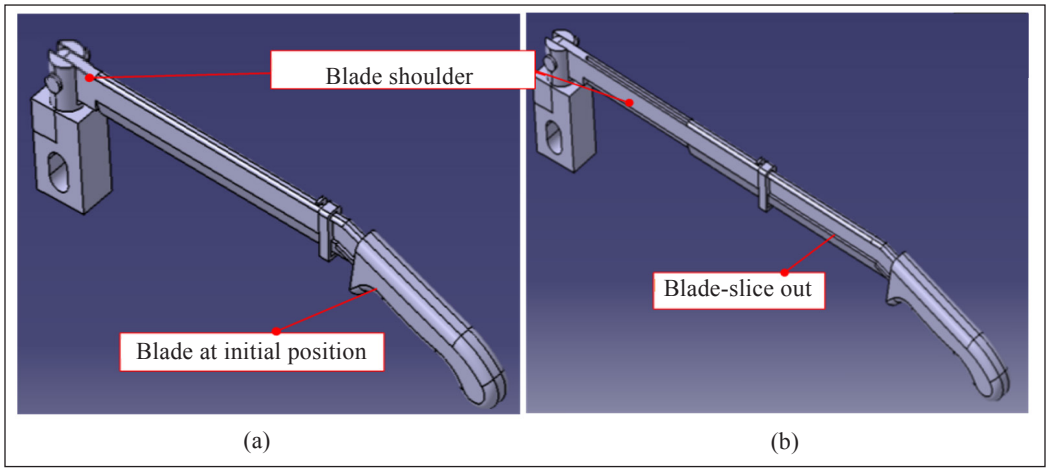


Figure 10. Slicing movement of the blade with respect to its holder

Blade Slicing Feature

In addressing the main problem (to eliminate the chopping process) of the development of the apparatus, the slicing method is proposed to replace the chopping technique. The slicing feature will ease the process of removing the husk by facilitating the blade to make the first cut on the husk. The process of removing the husk resembles the wood sawing process as shown in Figure 10.

Knife angle is the angle of the blade which is referred to the X-axis viewed from the XY plane shown in Figure 11. By putting the blade in an angled position will reduce the surface area contact between the husk and the respective blade area as justified by Pressure's law.

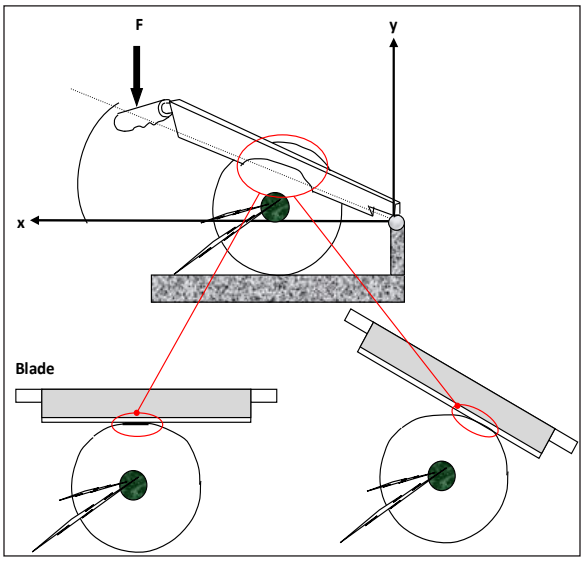


Figure 11. The blade angle during cutting process

Product Modeling

The young coconut processing apparatus was modeled using CATIA V5R19 as shown in Figure 12. The software is chosen because the application provided makes it easy to manipulate the model and thus, convenient to alter the 3D buildup. In product modeling, the solid design is used. All components and subassembly of the product are assembled as to resemble the final product.

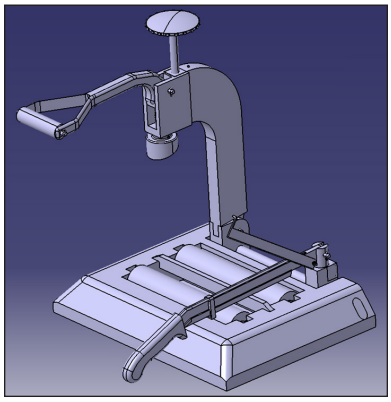


Figure 12. Young coconut processing apparatus modeling using CATIA V5R19

Punch Bit Head Analysis

The pressing force that acts on the roller of the punch lever is transmitted to the punch bit head. An amount of 1471N is required to penetrate the top endocarp of the young coconut fruit. Thus, the minimum pressing force to be applied to punch lever is equivalent to the penetrating force as shown in Figure 13.

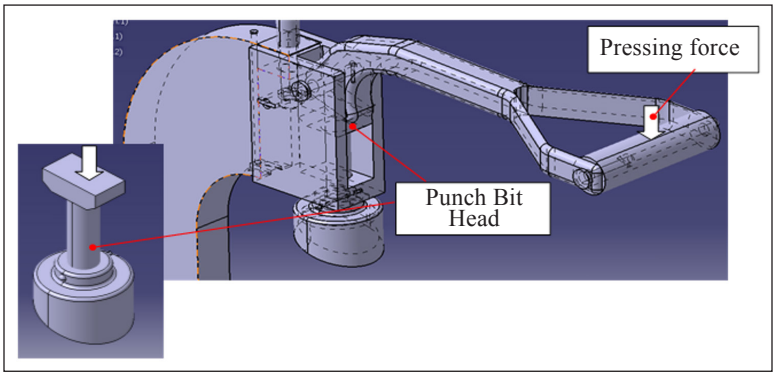


Figure 13. Punch bit head

The stress applied on this designed component were analyzed based on these two justifications; (i) if the applied stress, α_{apl}, is greater than the material compressive yield strength, α_y, and (ii) if the buckling force is smaller than the applied stress, α_{apl}. Thus, it can be said that, if one of the justification is true, then the component with the defined diameter is considered fail.

Consider that the selected material is cast iron ASTM A-197 with modulus elasticity of 172 GPa and its compressive ultimate strength, α_{ult}, is 572 MPa.

$$\sigma_{apl} = \frac{F}{A} = \frac{1471N}{3.1416 \times 10^{-4} m^2} = 4.6823MPa$$

Inculcating the safety factor of 1.5 into the α_{apl}. This is to defined that the system will not fail if under excessive force. Thus the new applied stress,

$$\sigma_{new} = safety\ factor \times \sigma_{apl} = 7.0235\ MPa$$

$$F_{new} = safety\ factor\ \times\ \sigma_{apl} = 2206.5\ N$$

It shows that, $\alpha_{apl} < \alpha_{ult}$, thus the punch bit head with diameter of 20 mm does not fail under the compressive stress of 7.0235 MPa.

Meanwhile, according to the Euler Buckling Column,

$$F_{buck} = \frac{\pi^2 EI}{(KL)^2} = 12.62\ MPa$$

Where E is the modulus of elasticity, I is the area moment of inertia, K is the column effective length factor, and L is length.

With the F_{buck} = 12.62 MPa and F_{new} = 2206.5 N,it shows that the $F_{buck} > F_{new}$, thus the round shaft with diameter of 20 mm won't buckle if a force of 2206.5 N is exerted on the shaft.

Based on the weight and the design proposed, the device is categorized as a portable and easy to handle. Although most of the components are suggested to be made of iron and steel, the total weight of the apparatus is 29.171 kg, where it became the lightest young coconut processing apparatus available in the market. This is finalized using the utilization of 3D modeling software CATIA V5R19 which instantly shows the effect on the weight by changing the material used.

CONCLUSION

Although the device is still not being materialized, but based on engineering analysis and the 3D product modeling, it shows that the potentials and the advantages of the device. In general, by executing a proper study on the development of any product, the design of the device can be optimize prior to the development process. It also can reduce tremendously the possibility of error to occur. The device developed eliminates the hazardous task by eliminating the chopping process and replacing it with the slicing process. The blade of the device is designed to allow the slicing movement to be maneuvered during the husk removal process. Therefore, the tendency to get caught in accidental injury during the chopping process can be significantly reduced. Meanwhile, the puncher has replaced the usage of the copper in creating an opening at the top of the endocarp. Thus, the opening can be made easily whereby it is not possible conventionally.

ACKNOWLEDGMENTS

The author would like to thank to the Ministry of Higher Education Malaysia and Universiti Putra Malaysia (UPM) for supporting this work. The author, Muhamad Hellmy Hussin would like to thank to Universiti Kuala Lumpur Malaysia France Institute (Unikl MFI) and Majlis Amanah Rakyat (MARA) for granting a fellowship for his Masters study.

REFERENCES

Abraham, A., Girish, M., Vitala, H. R., & Praveen, M. P. (2014). Design of harvesting mechanism for advanced remote-controlled coconut harvesting robot (A.R.C.H-1). *Indian Journal of Science and Technology 7*, 1465-1470.

Adzimah, S. K., & Turkson, S. O. (2015). Conceptual design of coconut dehusking machine. *Innovative Systems Design and Engineering 6*, 1-10.

Ismail, N. F., Sulong, A. B., Muhamad, N., Tholibon, D., Md Radzi, M. K. F., & Wan Ibrahim, W. A. S. (2015). Review of the compression moulding of natural fiber-reinforced thermoset composites: Material processing and characterisations. *Pertanika Journal Tropical Agricultural Science, 38*(4), 533-547.

Jarimopas, B., Ruttanadat, N., & Terdwongwarakul, A. (2009). An automatic trimming machine for young coconut fruit. *Biosystems Engineering 103*, 167-175.

Ketan, K. T., Vinod, P. S., & Sakhale, C. N. (2014). Design and performances of coconut de-shelling machine. *International Journal of Engineering Research and Applications, 4*, 39-44.

Lihua, W. (2015). Design of coconut feeding machine based on motion controller. *Journal of Applied Science and Engineering Innovation 2*, 160-162.

Mani, A., & Jothilingam, A. (2014). Design and fabrication of coconut harvesting robot: COCOBOT. *International Journal of Innovation Research in Science, Engineering and Technology 3*, 1351-1354.

Mohanraj, A. P., Rahgul Krishna, S., Kannan, S. S., Rajkumar, M., & Elango, A. (2014). Design and analysis of a coconut harvesting maneuver. *International Journal of Scientific & Engineering Research 5*, 508-512.

Prashant, Y., Gopinath, C., & Vignesh, R. (2014). Design and development of coconut fiber extraction machine. *SasTech Journal 13*, 64-72.

Sabale, R., & Kolhe, K. P. (2015). Review of coconut fiber extraction machines. *International Journal of Science, Engineering and Technology Research 4*, 1269-1273.

Salleh, J., Mohd Yusoh, M. K., & Ruznan, W. S. (2015). Tensile strength of some natural-fibre composites. *Pertanika Journal Tropical Agricultural Science, 38*(4), 575-582.

Sangameshwara, G. S., & Ravaikiran, B. S. (2015). Design and fabrication of portable virgin coconut oil extracting machine. *International Journal of Engineering Research and General Science, 3*, 271-275.

Satip, R., & Kiattisak, R. (2008). Design and development of semi-automatic cutting machine for young coconuts. *Maejo International Journal of Science and Technology*, 1-6.

Sim, F. S., Mohd Irwan Lu, N. A. L., Lee, Z. E. T., & Mohamed, M. (2015). Removal of dissolved organic carbon from peat swamp runoff using assorted tropical agriculture biomass. *Pertanika Journal Science & Technology, 23*(1), 51-59.

Syafriani, R., Sukandar, E.Y., T., Apriantono., & Sigit, J. I. (2014). The effect of coconut water (*Cocos nucifera L.*) and an isotonic drink on the change of heart rate frequency in the rats induced hypertension. *Procedia Chemistry, 13*, 177-180.

Vasconcelos, B., & Junior, B. B. (2015). The causes of work place accidents and their relation to construction equipment design. *Procedia Manufacturing, 3*, 4392-4399.

Venkataramanan, S., Abhinav Ram, B., & Rahul, R. (2014). Design and development of automated coconut dehusking and crown removal machine. *Internationa; Journal of Sciences: Basic and Applied Research, 13*, 183-219.

Vidhan Sigh, T., & Udhayakumar, R. (2013). Development of a coconut de-shelling machine. *Tropical Agricultural Research, 18*, 1-8.

Vinod, P. S., Ketan, K. P., & Chandrashekhar, N. S. (2014). Design and development of coconut de-husking machine. *International Journal of Engineering Research and Technology 3*, 1-6.

Yahya, S., & Zainal, M. (2014). Design and performance of young coconut shaping machine. *Journal of Tropical Agricultural and Field Sciences 42*, 19-28.

Yuniwarti, E. Y. W., Asmara, W., Artama, W. T., & Tabbu, C. R. (2015). Virgin coconut oil supplementation increased the survival of Avian Influenza Virus (H5N1) infected chicken. *Asian Journal of Poultry Science*, *9*(2), 106-111.

Pertanika J. Sci. & Technol. 26 (4): 2045 - 2062 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

The Adsorption of Reactive Orange 16 and Basic Blue 3 from Aqueous Solution Using Quaternized Blue Swimmer Crab Carapace

Y. J. Chuah., Y. P. Tan* and A. H. Abdullah

Department of Chemistry, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ABSTRACT

The crab carapace is a waste which cannot be decomposed. This waste was used to remove the Reactive Orange 16 (RO16) and Basic Blue 3 (BB3) from aqueous solution at different operational parameters such as pH, mass load, the concentrations of dye and the temperature. The crab collected was modified to obtain quaternized crab (QC) using (3-chloro-2-hydroxypropyl) trimethylammonium chloride solution ($C_6H_{15}Cl_2NO$, 65% w/w in water). The pH of the dyes solution was varied from pH 4 to 10. The highest adsorption percentage was achieved at pH 7 for both dyes. Increasing the QC mass for the adsorption process had granted an increase of dyes removal percentage. The highest adsorption percentage was achieved at 91.00% for RO16 and 29.40% for BB3 dyes with 7.5 g/L QC used. However, the adsorption capacity of QC decreased with higher QC mass because the dye molecules occupied on the surface and prevented other molecules to diffuse into the QC. At higher concentration beyond 20 mg/L and 10 mg/L of RO16 and BB3, respectively, the maximum adsorption was achieved at 2.5362 mg/g and 0.6812 mg/g. The adsorption of both dyes by QC was best fitted using Langmuir isotherm model, explaining the adsorption mainly occurred as a single layer on the surface of QC. Comparison to the results obtained from the kinetic models, the adsorption was chemisorption in nature. According to the thermodynamic studies, the adsorption of RO16 was an exothermic, while BB3 was an endorthermic process.

Keywords: Adsorption, BB3, crab, RO16

ARTICLE INFO

Article history:
Received: 13 November 2017
Accepted: 09 April 2018
Published: October 2018

E-mail addresses:
typ@upm.edu.my (Y. P. Tan)
jason_youjian@outlook.com (Y. J. Chuah)
halim@upm.edu.my (A. H. Abdullah)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Water as one of the renewable resources becomes an environmental issue when the percentage of usable clean water is decreasing. The contaminants found in the water proved that the wastewater from various industries has not been treated before discharging and the pollutants spread to various water resources. Textile industry uses more than 8000 chemicals for the manufacture processes such as dyeing and printing. A study by Kant (2012) stated that among 72 toxic chemicals used in textile industry, 30 of which could not be removed. This explains why the textile manufacturing is a cause of environmental problems.

Dyes are commonly found in the wastewater especially the untreated wastewater discharged from factories. Dyes are widely used in industries such as textile industry, paper industry, wood or lodging industry, furniture industry, stationery industry and others. Dyes can be grouped into different classes such as cationic dyes, anionic dyes, acid dyes, azo dyes, dispersive dyes, direct dyes and so on. Different classes of dyes are applied onto different products. Sometimes, there are combinations of different classes of dyes which can be applied onto a product. The dyes which are dissolved in water and discharged from the industries will flow into water streams. Some dyes are non-biodegradable and the existing treatments such as precipitation, coagulation and flocculation, ion exchange and membrane separation were not able to remove all dyes from the water. Those dyes are mostly synthetic dyes which are harmful to living things. Hence, it is crucial to remove the dyes from wastewater before releasing to the environment.

Adsorption is one of the efficient methods to remove pollutants from wastewater. It is sometimes being used in the water treatment together with other methods since it has the ability to remove certain contaminants from wastewater. Besides, it is a low-cost method, easy to operate and insensitivity to harmful materials and toxic substances (Chatterjee et al., 2012). Adsorption technology is conducted by adding the adsorbent into the wastewater to remove the contaminants from the water. Adsorption occurs on the surface of the solids (adsorbent) by weak van der Waal's forces (physisorption), covalent bonding (chemisorption) and electrostatic forces. The adsorbent used can be the solid wastes we generate such as fruit peels depending on the adsorption capacity of the materials, hence there is an opportunity to reduce the solid wastes produced since those wastes can be transformed into useful materials. The wastes to be used as adsorbents must fulfil the requirements, where low generation cost, easy to be obtained, and will not introduce any harmful or toxic substances into the water.

Crabs are decapod crustacean, living under the sea, served for the many hence creating a large amount of waste such as carapace. Crab carapace is the upper section of the shell of crab. It contains about 30% of chitin, the rest of the contents are mostly protein (11-29%), calcium carbonate (40-66%), and a small portion of lipids (Kim, 2003; Rae et al., 2009; Lu et al., 2007). Some researchers used chitin extracted from crab shell as adsorbent in their work (Akkaya et al., 2007).

Chitosan is produced by treating the chitin using alkaline substance such as sodium hydroxide. Esquerdo et al. (2014) reported the effectiveness of chitosan obtained from shrimp waste to remove 5 food dyes which were FD&C blue 2, red 40, yellow 5, yellow 6 and Food red 2. Another research was carried out to compare the influence of different chitosan base on azo reactive dye (RR120) from aqueous solution. The chitosan nanoparticle shows higher adsorption capacity and faster adsorption kinetics compared to dissolved chitosan (Momenzadah et al., 2011).

Szygula et al. (2009) reported the anionic dye (Acid Blue 92) removal was achieved at about 99% using coagulation-flocculation method, but a larger amount of chitosan was required in higher pH solution. However, the production of sludge during the process was low.

In the present work, crab carapace was used as the adsorbent for the entire work to test on the effectiveness in the removal of cationic and anionic dyes. The crab carapace is a waste and will be discarded as it cannot be consumed. The generation of this waste without treated caused an environmental issue since it cannot be decomposed. The cationic dye used in the work was Basic Blue 3 (BB3) and the anionic dye was Reactive Orange 16 (RO16). The adsorption of RO16 and BB3 by crab carapace was investigated in batch system to study the influence of operational parameters such as the pH of dyes solutions, the adsorbent mass, the concentrations of the dyes and the temperature.

MATERIALS AND METHODS

Chemicals and Materials

In this study, the chemicals were used without purification. (3-chloro-2-hydroxypropyl) trimethylammonium chloride solution, 65% w/w in water, RO16 and BB3 dyes were purchased from Sigma Aldrich. Portunus pelagicus or simply blue swimmer crabs were collected at Pangkor Island, Malaysia.

The carapaces of blue swimmer crab were washed with distilled water and soaked for a day. The sample was then boiled for an hour to soften the skin. Subsequently, it was washed to remove the residue with distilled water and dried in oven at 80°C overnight. The steps were repeated until the soft skin from the inner part of the carapaces was fully removed. The dried carapaces were blended and ground using blender, mortar and pestle to obtain powdered raw crab carapace (RC). The powdered sample was then sieved to pass through 1 mm sieve to ensure the particle size of the sample was less than 1 mm before storing in a desiccator.

The quaternization of the sample was conducted using (3-chloro-2-hydroxypropyl) trimethylammonium chloride solution ($C_6H_{15}Cl_2NO$, 65% w/w in water) which was adjusted to pH 8 using 5 M NaOH solution (Fan et al., 2012). 15 g of the powdered raw crab was soaked in 115 mL of the $C_6H_{15}Cl_2NO$ solution and heated at 50°C for 24 hours in the oven

with intermittent stirring. The mixture was rinsed with distilled water and dried overnight in the oven at 50°C. The sample was labelled as QC.

Characterization of RC and QC

The functional group vibrations were elucidated using Fourier-Transform Infrared (FTIR) spectrometer (Perkin-Elmer FTIR, 1725 X). The RC and QC were dried and mixed with KBr. It was then pressed with the aid of a bench press and the IR spectra were recorded. The surface morphology of the RC and QC were studied using scanning electron microscopy (SEM) method by energy dispersive X-ray spectrometer (SEM-EDX)-JEOL JSM-6400. The RC and QC were dried and put on top of the stubs which was then coated with a thin layer of gold to prevent the charging effect.

Dyes Adsorption (Batch System)

The adsorption of RO16 and BB3 from aqueous solution was conducted in a batch system to study the ability of QC on adsorbing the dye molecules with different operational parameters such as the initial pH, QC mass, initial concentration and temperature. The experiments were done using orbital shaker at 150 revolutions per minute (rpm) at room temperature. The initial pH of the dye solution was controlled at the range of pH 4 to 10 using NaOH and HCl. A preliminary test was carried out to determine the range of QC mass used. The adsorption equilibrium was reached at 7.5 g/L. Hence the QC mass used was varied from 2.5 to 7.5 g/L.

The adsorption isotherm was determined by conducting an experiment agitating 20 mL RO16 and BB3 dye solutions. The concentration of RO16 was varied from 10 to 30 mg/L and for BB3, the concentration was altered in the range of 5 to 15 mg/L. The effect of temperature of dyes adsorption was studied from 30 to 70°C.

The residual dye concentrations were determined using UV-Vis spectrophotometer (Shidmazu UV-1650 PC) at λ_{max} of 494 nm for RO16 and 654 nm for BB3. The removal percentage of RO16 and BB3 and adsorption capacity of the QC was calculated using the equations below:

$$\text{Removal percentage, \%} = \frac{C_o - C_e}{C_o} \times 100 \quad (1)$$

$$\text{Adsorption capacity, } q_e = (C_o - C_e) \times \frac{V}{m} \quad (2)$$

where C_o and C_e are the concentrations (mg/L) of the dye solutions before and after the experiments, V is the volume of the dye solution (L) and m is the mass of QC used (g).

RESULTS AND DISCUSSION

Characterization of RC and QC

Fourier-Transform Infrared Spectroscopy (FTIR). The adsorption ability of the adsorbent is relied on the functional groups. The surface functional groups of RC and QC are illustrated in the FTIR spectra shown in Figure 1. The broad band region of 3200 cm^{-1} observed is due to the stretching of OH and N-H vibration. The infrared spectra of both raw crab (RC) and quaternized crab (QC) show similar spectra, it is believed that there is overlapping between the OH and N-H stretching in the same absorption region. The OH stretching in QC is lower because most OH groups on the RC were replaced. At 1650 cm^{-1}, both infrared spectra of RC and QC show no significant difference because the functional groups of C=O vibration of acetyl groups can be found in RC and QC. However, at region 1000 cm^{-1} to 1500 cm^{-1}, the functional groups for both RC and QC are the same, but broader band for QC in the spectrum. The peak at 1400 cm^{-1} is resulted from the combination of N-H deformation and C-N stretching vibration. The absorbance intensity of infrared spectrum in QC is higher than RC. This shows an increase in C-N bonding in the QC. At 1154 cm^{-1}, it is due to the anti-symmetric stretching of the C-O-C bridge, which can be found in chitosan (Peniche et al., 1999). The similarity in RC and QC spectra might be due to the limitation of the sensitivity of the spectrometer as the quaternization process would alter the functional groups of RC.

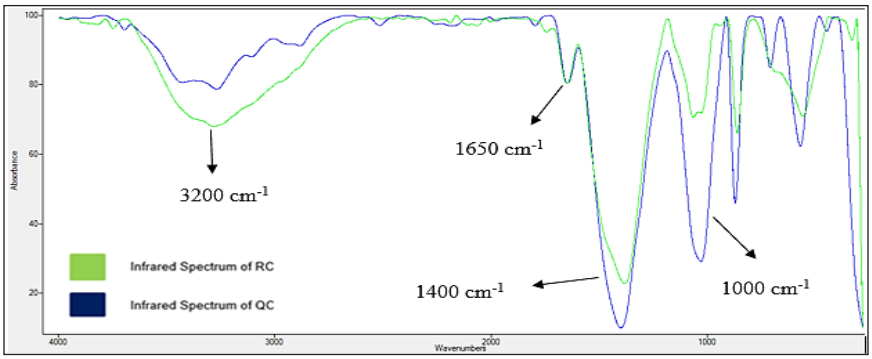


Figure 1. FTIR spectra of RC and QC

Scanning Electron Microscopy (SEM)

The morphological and surface characteristics of QC was examined by scanning electron microscopy (SEM). Figure 2 shows a micrograph of QC captured at magnification of 1000 times. The surface of QC is lumpy and cluster which helps in adsorbing the dye molecules on the surface of QC (Subramani and Thinakaran, 2017). Mohamed et al. (2015) proposed that the surface of QC was covered by methyl groups after quaternization.

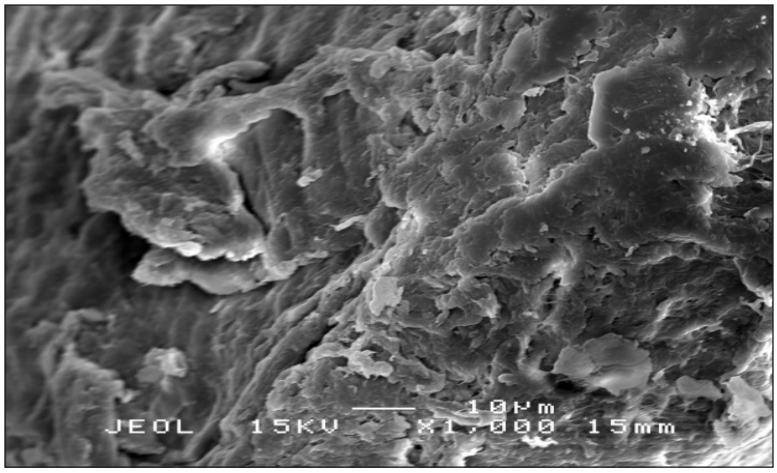


Figure 2. SEM micrograph of QC

Batch Studies of RO16 and BB3

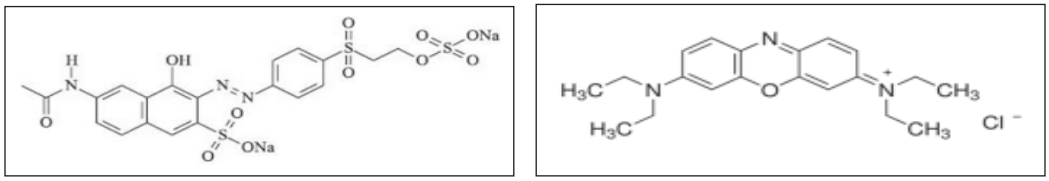
Effect of pH. The pH of dye solutions plays an important role in experiments since it affects the ability of QC to adsorb the dye molecules from aqueous solution. The pH of the solutions takes part in improving the surface charge of the QC or vice versa. At low pH, the adsorption of BB3 must be low but achieves high removal percentage of RO16 theoretically. It is because more protons are generated at low pH and lead to the protonation of the amine groups on QC to form $-NH_3^+$ (Chiou et al., 2004). The RO16 adsorption by QC at low pH with the existence of $-NH_3^+$ is high due to the electrostatic attraction between the oppositely charges of RO16 and QC. Increasing the pH of RO16 would decrease the removal percentage because more negatively-charged sites appear after the deprotonation of the QC. An increase in pH of BB3 increases the removal percentage because the negatively-charged sites on the QC are more favor to attract the BB3 dye molecules. The chemical structures of RO16 and BB3 are shown in Figure 3. Figure 4 shows the effect of pH on the adsorption of RO16 and BB3. The removal percentage of BB3 at pH 8 is lower than pH 7. A report by Rosa et al. (2008) had shown that the adsorption of some dyes was independent to the pH of the dye solution. It is believed that the adsorption reaction on QC has insufficient reactive sites at certain pH. The ratio of the reactive sites on the surface of QC to the amount of dye molecules is another factor to consider during the adsorption reaction. For RO16 adsorption, the ability of QC to adsorb the dye molecules is weaker at high pH due to the repulsion between the RO16 dye molecules with the negatively-charged surface group. According to Chiou et al. (2002), decreasing the pH of the anionic dyes would lead to an increase in removal percentage. However, at lower pH than pH 7, excess anions appear to compete for adsorption and limit the uptake of the dye molecules (Guibal, 2004). In addition, not all reactive amine sites are responsible to adsorb the dye molecules.

There is a chance that the reactive sites can form inter and intramolecular hydrogen bonds which further limit the adsorption (Crini et al., 2008). The adsorption mechanism of BB3 and RO16 dye molecules are proposed as below:

$$R - CH_2 - O^- H^+ + BB^+ \rightarrow R - CH_2 - O^- BB^+ + H^+$$

$$R - N^+(CH_3)_3 + RO16 - SO_3^- \rightarrow R - N(CH_3)_3 - O_3S - RO16$$

where R represents crab carapace, BB^+ and $RO16 - SO_3^-$ represent BB3 and RO16 molecules, respectively.



(a) (b)

Figure 3. Chemical structures of (a) RO16 and (b) BB3

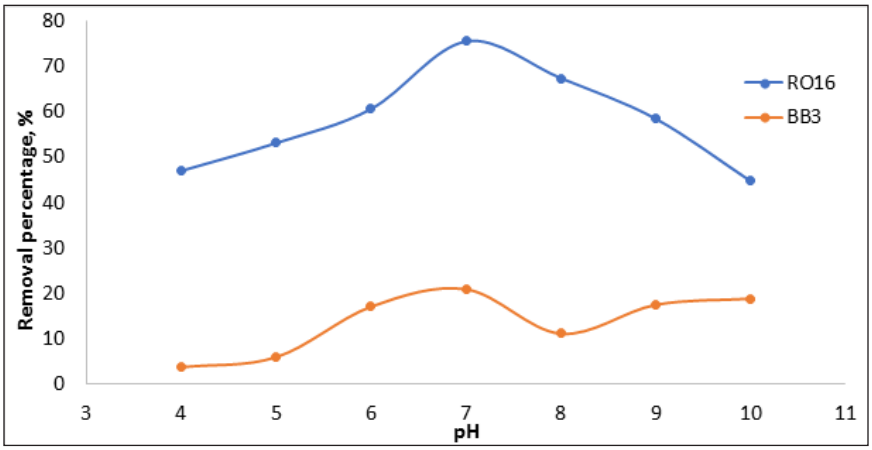


Figure 4. Effect of pH on RO16 and BB3 adsorption

Effect of QC Mass. The effect of increasing QC mass on the dyes removal percentage and adsorption capacity is depicted in Figure 5. An increase in QC mass increases the removal percentage of RO16 and BB3. The increase of QC mass offers more surface area and reactive sites for adsorption (Vijayaraghavan et al., 2006). The adsorption capacity is determined by calculating the amount of dyes adsorbed per unit weight of the QC.

Increasing QC mass in the experiment caused a decrease in adsorption capacity because less dye molecules were adsorbed per unit weight. The overlapping of the adsorption sites could be another reason leading to the decrease of the adsorption capacity with increasing QC mass (Subramani & Thinakaran, 2017). The diffusion of dye molecules onto the reactive sites of QC are harder due to the overlapping of the sites (Crini et al., 2008).

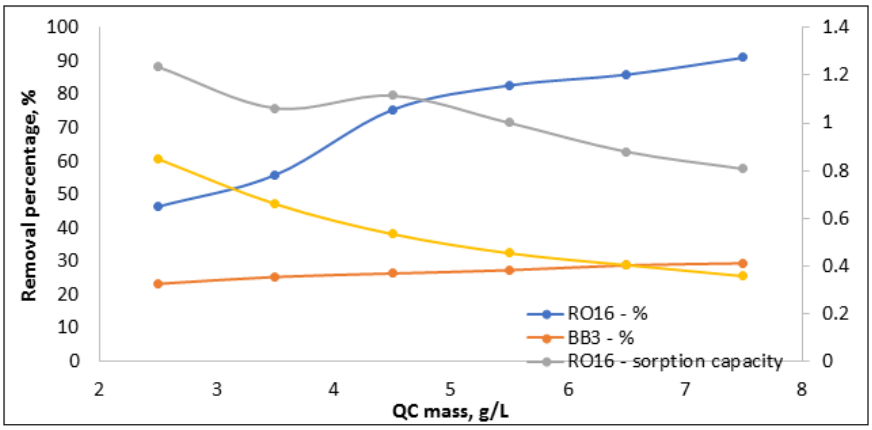


Figure 5. Effect of the QC mass on RO16 and BB3 adsorption

Effect of Dye Concentration. Figure 6 shows the removal percentage of RO16 and BB3 by QC. From Figure 6, at higher concentration, the adsorption capacity of RO16 and BB3 reduced. According to Malik (2003), there are a few steps for the dye molecules to be adsorbed onto the reactive sites of the adsorbent. The dye molecules will first overcome the boundary-layer effect and diffuse into the surface of the adsorbent. The last step is entering the porous sites of the adsorbent. From Figure 6, the adsorption capacity for both dyes increased at lower concentration. However, it is decreased at higher concentration beyond 10 mg/L for BB3 and 20 mg/L for RO16. This is due to the maximum adsorption capacity has been reached at those concentrations.

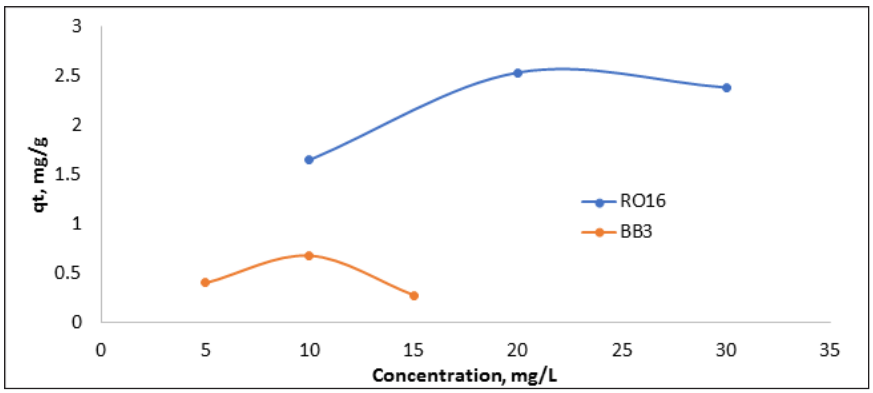


Figure 6. Adsorption capacity of QC on removal of RO16 and BB3 at different concentrations

Effect of Temperature. From Figure 7, the removal percentage of BB3 increased from 30 to 70ºC. At high temperature, the diffusion rate of the dye molecules onto the adsorbent is faster. It is believed that the swelling effect of the QC at high temperature enables some larger dye molecules to penetrate further (Mane et al., 2007) which leads to a higher removal percentage. Increasing the temperature of the dye solutions will increase the removal percentage but reduces the adsorption capacity (Crini et al., 2008). When the temperature of the dye solution increases, the solubility of the dye also increases. The interaction forces between solute and solvent become stronger compared to those between solute and adsorbent. Hence, the solute is more difficult to adsorb. However, a report proposed the dye adsorption was not affected by the temperature especially for anionic dyes. Dye adsorption is an exothermic process which will weaken the bonds between the adsorbent and the dye molecules, hence, there is a controversy (Kumar, 2000).

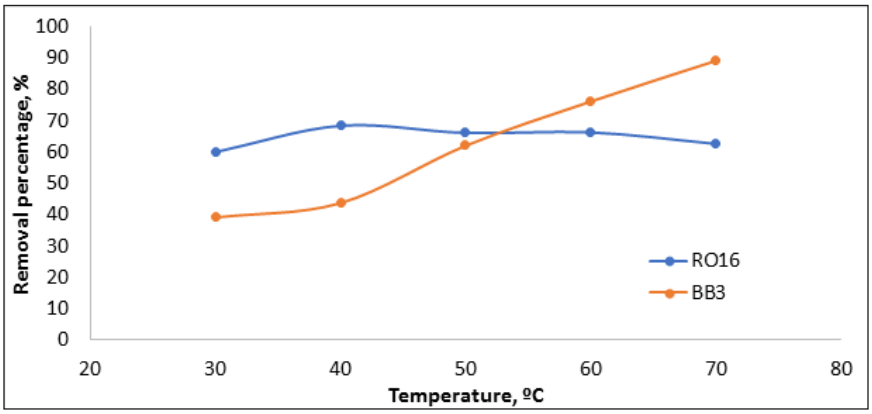


Figure 7. Effect of temperature for the RO16 and BB3 adsorption

The RO16 adsorption did not increase after 40ºC using QC. This is because the bonds between the dye molecules and the reactive sites would be weakened with increasing temperature. This leads to a low adsorption ability. At high temperature, the solubility of dye increases, the QC will be difficult to adsorb the dye molecules because the dye molecules keen to bond with the solute rather than QC (Crini et al., 2008).

Isotherm Modelling. The Langmuir isotherm suggests that the adsorption takes place at specific sites of the adsorbent. The reactive sites of the adsorbent are all equivalent where the dye molecules are able to bind at the sites without being affected by nearby sites. Each reactive site is only occupied by a dye molecule and there is no further adsorption process taking place at the same site. The linearized form of Langmuir isotherm equation used in the investigation is as below (Rosa et al., 2008):

$$\frac{C_e}{q_e} = \frac{1}{K_L q_m} + \frac{1}{q_m} C_e \quad (3)$$

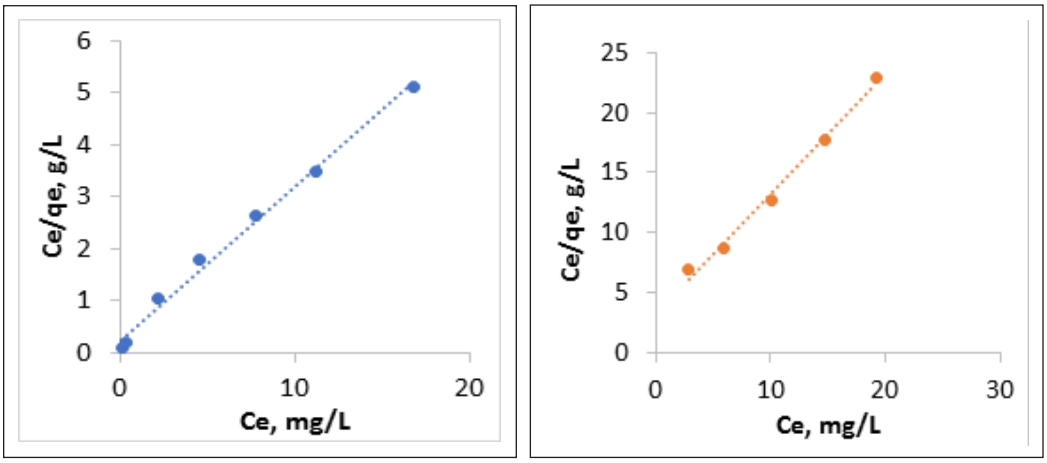
where C_e is the equilibrium concentration of the dye (mg/L), q_e is the amount of dye adsorbed per unit weight of adsorbent at equilibrium (mg/g), q_m is the maximum adsorption capacity (mg/g) and K_L is the adsorption equilibrium constant (L/mg).

The adsorption isotherm can be investigated not only using Langmuir model but Freundlich model which is famous in the adsorption process to determine whether the adsorption of the sorbate on the adsorbent occurs at multilayer. The Freundlich isotherm is used on heterogeneous surfaces during the adsorption between the adsorbent and the dye molecules. The linear form of Freundlich isotherm equation is represented by:

$$\ln q_e = \ln K_F + \frac{1}{n_F} \ln C_e \quad (4)$$

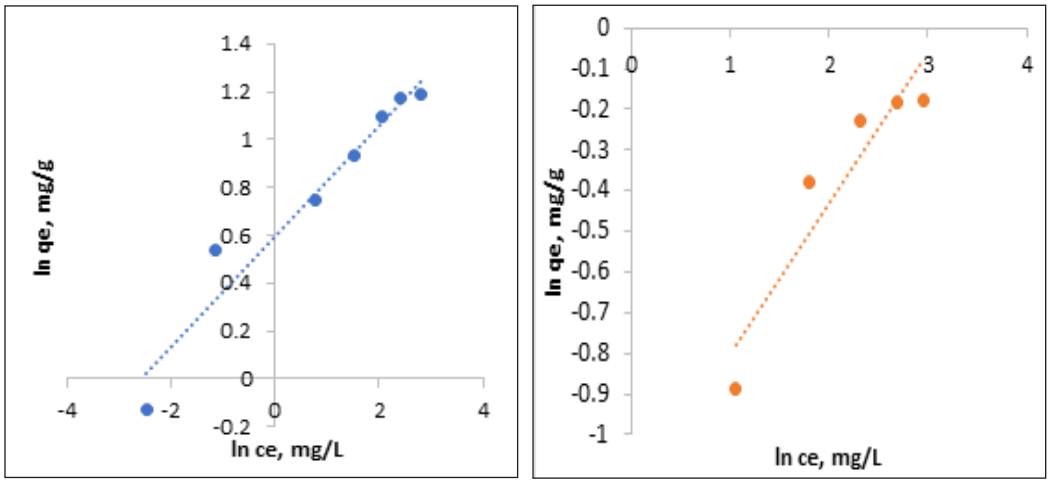
where q_e is the amount of dye adsorbed per unit weight of adsorbent at equilibrium (mg/g), C_e is the equilibrium concentration of the dye (mg/L), K_F is the Freundlich constant for adsorption capacity and n_F is the Freundlich constant for adsorption intensity.

Figure 8 shows the relationship between the ratio of dye molecules adsorbed on QC and the remaining concentration. The Langmuir constants and the maximum adsorption capacity of QC can be obtained from the graphs. The linear graphs of the plots show that the adsorption of both dyes on QC happened on the homogeneous surface besides forming a monolayer adsorption. The high correlation coefficient of the plot best describes the reaction takes place only at single layer on the adsorbent surface and with constant activation energy (Crini et al., 2008). Figure 9 shows the Freundlich isotherm for the adsorption of RO16 and BB3. The plots show straight lines, however with low correlation coefficient obtained for the adsorption. The Freundlich constants and the correlation coefficients of both dyes were shown in Table 1. The low correlation coefficient shows the adsorption is not a multilayer adsorption process. However, the comparison between both isotherms explaining that the adsorption of both dyes is best fitted using Langmuir isotherm with higher correlation coefficient.



(a) (b)

Figure 8. Langmuir linear fit for: (a) RO16; and (b) BB3 adsorption



(a) (b)

Figure 9. Fruendlich fit of isotherm for: (a) RO16; and (b) BB3 adsorption

Table 1
Langmuir and Freundlich Constants for the RO16 and BB3 Adsorption by QC

	Langmuir isotherm			Freundlich isotherm		
Dyes	q_m (mg/g)	K_L (L/mg)	R^2	n_F	K_F	R^2
RO16	3.3887	1.2514	0.9936	1.8090	4.3029	0.9468
BB3	1.0031	0.3047	0.9899	0.3110	2.7122	0.8697

Kinetics Studies

The consideration for the use of an adsorbent to remove unwanted pollutants is not only the adsorption capacity of the adsorbent, the rate of adsorption plays an important role in water treatment. The adsorption kinetics must be determined to find out the rate of the adsorption and the factors involved in the reactions. There are three kinetic models used to explain the rate of the adsorption process and adsorption mechanism (Crini et al., 2008). These three models are (i) Lagergren model, (ii) Ho and McKay model, and (iii) Weber and Morris model. The models are best used to explain the adsorption on surface of the adsorbent, chemical reactions involved in the process and the effect of diffusion through the boundary layer (Annadurai et al., 2008).

Lagergren model with pseudo-first-order equation is used to investigate the adsorption of dyes. The linear form of pseudo-first-order equation is given by:

$$\log(q_e - q_t) = \log q_e - \frac{k_1}{2.303} t \quad (5)$$

Ho and McKay equation with pseudo-second-order equation has been used to explain the adsorption kinetic of an adsorbent in removing the pollutants. According to Cochrane et al. (2006), this equation can be used to explain the rate-limiting process which is due to the chemisorption reaction. There is a sharing or exchanging of electrons during the adsorption process. The linearized equation is given by:

$$\frac{t}{q_t} = \frac{1}{k_2 {q_e}^2} + \frac{1}{q_e} t \quad (6)$$

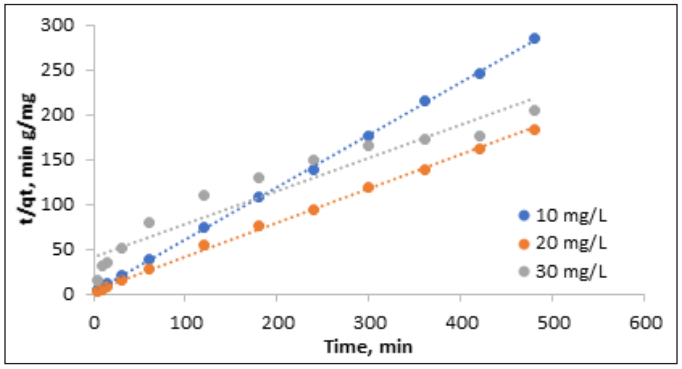
The intraparticle diffusion of the adsorption can be studied using Weber-Morris plot. The equation is represented by:

$$q_t = k_d t^{\frac{1}{2}} \quad (7)$$

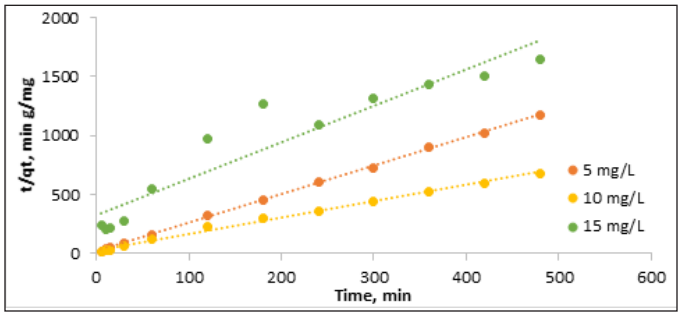
where q_e and q_t = amount of dye adsorbed per unit weight of adsorbent at equilibrium and at a certain time (mg/g), k_1 = the first-order-rate constant (min^{-1}), k_2 is the rate constant of pseudo-second-order adsorption (g/mg min) and k_d is the intraparticle diffusion rate constant (mg/g $min^{1/2}$).

Figure 10 shows the pseudo-second-order kinetic plots of BB3 dye using QC at different initial concentrations. From the graph plotted, the experiment data for 5 mg/L and 10 mg/L of the dye solution is best fitted with the pseudo-second-order kinetics. However, increasing the dye concentration to 15 mg/L, the correlation coefficient (R^2) is not achieved as high as 0.9900. According to Crini et al. (2008), increasing the dye concentration, the adsorption capacity will decrease as the surface reactive sites are fully

occupied, and the diffusion occurs through the boundary layer. The same phenomenon happened on the 30 mg/L of RO16 dye. The correlation coefficient (R^2) recorded at 0.9243 which is not good to explain the chemisorption mechanism of the adsorbent at this high dye concentration. The intraparticle diffusion rate constants for both BB3 and RO16 dyes in different concentrations can be obtained from the plots shown in Figure 11. From the plots, the intraparticle diffusion is not the main factor of kinetic process since the plots do not pass through the origin (Rosa et al., 2008), and the intraparticle diffusion is not the rate-determining step (Cochrane et al., 2006). According to Cochrane et al. (2006), the pseudo-second order model gives higher R^2 compared to intraparticle diffusion. This explains the rate-limiting step is a chemisorption which occurs between the dye molecules and the adsorbent. Moreover, since the plots do not pass through the origin, there might be some other processes controlling the rate of the adsorption (Crini et al., 2008). From the figures, there are three separate regions which can be observed. The first stage is the bulk diffusion with boundary layer effect, the next stage would be the intraparticle diffusion and the last stage is the final equilibrium stage, where the intraparticle diffusion starts to slow down (Mane et al., 2007). Table 2 shows a summary of rate constants and the correlation coefficients for the three models.

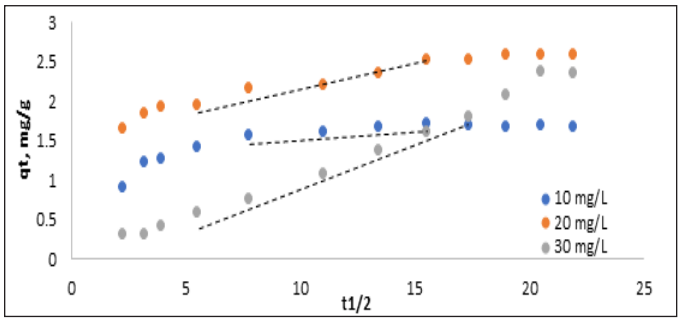


(a)

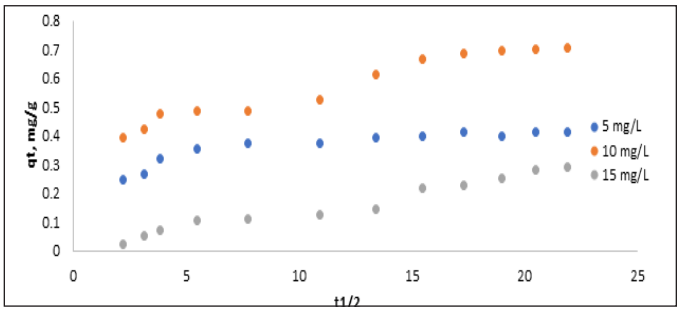


(b)

Figure 10. Pseudo-second-order kinetics of (a) RO16 and (b) BB3



(a)



(b)

Figure 11. Intraparticle diffusion of (a) RO16 and (b) BB3

Table 2
Kinetic parameters for the RO16 and BB3 Adsorption

Model	Parameters	RO16			BB3		
		Initial concentration (mg/L)					
		10	20	30	5	10	15
Pseudo-first-order	q_e exp (mg/g)	1.6732	2.5362	2.3840	0.3928	0.6686	0.2916
	q_e cal (mg/g)	0.5531	0.7386	2.1994	0.1067	0.2584	0.2826
	k_1 (1/min)	0.0210	0.0078	0.0048	0.0182	0.0074	0.0006
	R^2	0.8944	0.9349	0.9688	0.7480	0.8781	0.8970
Pseudo-second-order	q_e cal (mg/g)	1.7112	2.6469	2.7270	0.4145	0.7223	0.3236
	k_2 (g/mg min)	0.1317	0.0355	0.0032	0.3653	0.0769	0.0297
	R^2	0.9997	0.9985	0.9243	0.9995	0.9952	0.8981
Intraparticle diffusion	$k_d \times 10^{-2}$	1.95	5.14	10.53	0.45	2.41	0.43
	R^2	0.9733	0.9435	0.9964	0.9471	0.9517	0.9201

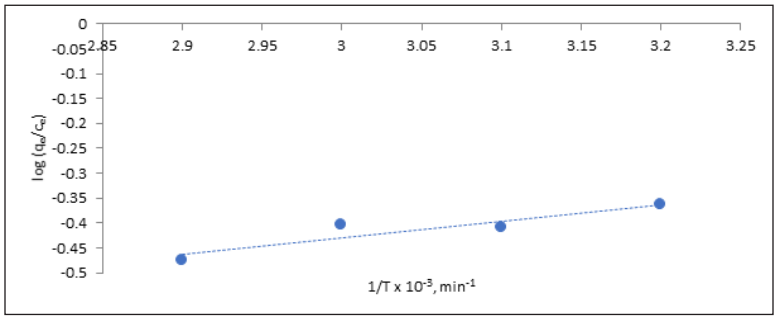
Thermodynamic Studies

The thermodynamic studies are important to determine what processes will occur spontaneously. The thermodynamic equation is shown below:

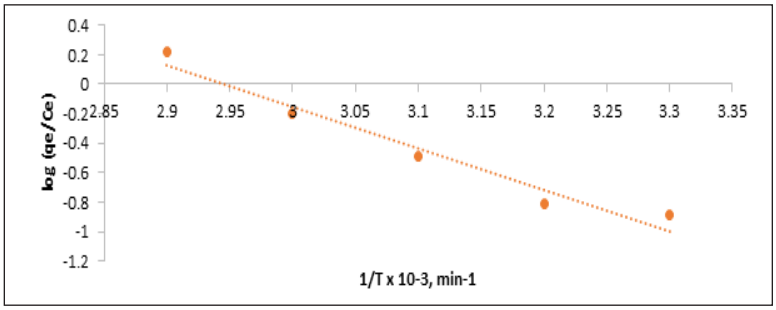
$$log\, K_d = \frac{\Delta S^0}{2.303\ R} - \frac{\Delta H^0}{2.303\ RT} \quad (8)$$

where K_d = distribution coefficient defined as q_e/C_e (L/g), q_e = amount of dye adsorbed per unit weight of adsorbent at equilibrium (mg/g), C_e = equilibrium concentration of dye remaining in solution (mg/L), T = absolute temperature (K). R = gas constant (8.3145 J/mol K), ΔS^0 = entropy change (J/mol K), ΔH^0 = enthalpy change (KJ/mol).

The Van't Hoff plots were shown in Figure 12, the slopes and the intercepts of the graphs were determined to obtain the values of enthalpy (ΔH^0) and entropy (ΔS^0). The sign of ΔH^0 and ΔS^0 can be positive or negative depends on the adsorption condition. A positive value of ΔH^0 indicates that the process is endothermic, while negative is exothermic.



(a)



(b)

Figure 12. Van't Hoff plots for (a) RO16 and (b) BB3 adsorption

Table 3
Enthalpy (ΔH°) and entropy (ΔS°) change

Dye	ΔH° (KJ/mol)	ΔS° (J/mol K)
RO16	-6.332	-27.219
BB3	54.222	159.810

From Table 3, the ΔH^{0} for BB3 adsorbed by QC were positive, showing the adsorption process was endothermic while the adsorption of RO16 was an exothermic process. The positive value of ΔS^{0} for BB3 dye indicates that the adsorption increased the randomness at the solid/solution interface. The negative value of ΔS^{0} for RO16 was preferable desorption rather than adsorption (Crini et al., 2008).

CONCLUSION

The quaternized crab (QC) was used for the adsorption of RO16 and BB3. QC showed a stronger adsorption of RO16 compared to BB3. The maximum adsorption of RO16 and BB3 was achieved at pH 7. This can be concluded as the adsorption process on removal of RO16 and BB3 was insignificant over the pH range of the dye solutions. Increasing the QC mass increased the removal percentage of RO16 and BB3 but decreased the adsorption capacity. This is due to the overlapping of the dye molecules on the surface of QC leads to a harder diffusion during the adsorption process. An increase on the concentration of dyes leads to a lower adsorption capacity as the maximum adsorption capacity has been reached. At high temperature, QC has lower adsorption ability to remove RO16, where the solubility of RO16 increases at high temperature, and the dye molecules keen to bond with the solute rather than QC.

The adsorption of RO16 and BB3 occurred mostly on the surface of QC. The dyes did not form multilayers phase on QC. The adsorption process followed pseudo-second-order kinetics with higher correlation coefficient (R^2) compared to pseudo-first-order kinetics. The evaluation of thermodynamic parameters for RO16 ($\Delta H^{\circ} = -6.332$) KJ/mol and $\Delta S^{\circ} = -27.21$ J/mol K) shows the exothermic process. For BB3 ($\Delta H^{\circ} = 54.222$ KJ/mol and $\Delta S^{\circ} = 159.810$ J/mol K), the adsorption was an endothermic process and a spontaneous reaction. The crab carapace as a waste is a potential adsorbent for water treatment and deserves future research for its uses in removing other pollutants.

ACKNOWLEDGEMENT

The authors wish to thank Graduate Research Fellowship (GRF) and Department of Chemistry, Faculty of Science of Universiti Putra Malaysia.

REFERENCES

Akkaya, G., Uzun I., & Güzel, F. (2007). Kinetics of the adsorption of reactive dyes by chitin. *Dyes and Pigments, 73*(2), 168-177.

Annadurai, G. (2002). Adsorption of basic dye on strongly chelating polymer: Batch kinetics Studies. *Iranian Polymer Journal, 11*(4), 237-244.

Chatterjee, S., Kumar, A., Basu, S., & Dutta, S. (2012). Application of response surface methodology for methylene blue dye removal from aqueous solution using low cost adsorbent. *Chemical Engineering Journal, 181*, 289-299.

Chiou, M. S., & Li, H. Y. (2002). Equilibrium and kinetic modeling of adsorption of reactive dye on cross-linked chitosan beads. *Journal of Hazardous Materials, 93*(2), 233-248.

Chiou, M. S., Ho, P. Y., & Li, H. Y. (2004). Adsorption of anionic dyes in acid solutions using chemically cross-linked chitosan beads. *Dyes and Pigments, 60*(1), 60-84.

Cochrane, E. L., Lu, S., Gibb, S. W., & Villaescusa, I. (2006). A comparison of low-cost biosorbents and commercial sorbents for the removal of copper from aqueous media. *Journal of Hazardous Materials, 137*(1), 198-206.

Crini, G., & Badot, P. M. (2008). Application of chitosan, a natural aminopolysaccharide, for dye removal from aqueous solutions by adsorption processes using batch studies: A review of recent literature. *Progress in Polymer Science, 33*(4), 399-447.

Esquerdo, V. M., Cadaval, T. R. S., Dotto, G. L., & Pinto, L. A. A. (2014). Chitosan scaffold as an alternative adsorbent for the removal of hazardous food dyes from aqueous solutions. *Journal of Colloid and Interface Science, 424*, 7-15.

Fan, L., Wu, P., Zhang, J., Gao, S., Wang, L., Li, M., … & Nie, M. (2012). Synthesis and anticoagulant activity of the quaternary ammonium chitosan sulfates. *International Journal of Biological Macromolecules, 50*(1), 31-37.

Guibal, E. (2004). Interactions of metal ions with chitosan-based sorbents: A review. *Separation and Purification Technology, 38*(1), 43-74.

Kant, R. (2012). Textile dyeing industry an environmental hazard. *Natural Science, 4*(1), 22-26.

Kim, D. S. (2003). The removal by crab shell of mixed heavy metal ions in aqueous solution. *Bioresource Technology, 87*(3), 355-57.

Kumar, M. N. V. R. (2000). A review of chitin and chitosan applications. *Reactive and Functional Polymers, 46*(1), 1-27.

Lu, S., Gibb, S. W., & Cochrane, E. (2007). Effective removal of zinc ions from aqueous solutions using crab carapace biosorbent. *Journal of Hazardous Materials, 149*(1), 208-217.

Malik, P. K. (2003). Use of activated carbons prepared from sawdust and rice-husk for adsorption of acid dyes: A case study of acid yellow 36. *Dyes and Pigments, 56*(3), 239-249.

Mane, V. S., Mall, I. D., & Srivastava, V. C. (2007). Use of bagasse fly ash as an adsorbent for the removal of brilliant green dye from aqueous solution. *Dyes and Pigments, 73*(3), 269-278.

Mohamed, R. R., Elella, M. H., & Sabaa, M. W. (2015). Synthesis, characterization and applications of N-quaternized chitosan/poly(vinyl alcohol) hydrogels. *International Journal of Biological Macromolecules, 80*, 149-161.

Momenzadeh, H., Tehrani-Bagha, A. R., Khosravi, A., Gharanjig, K., & Holmberg, K. (2011). Reactive dye removal from wastewater using a chitosan nanodispersion. *Desalination, 271*(1-3), 225-230.

Peniche, C., Argüelles-Monal., Davidenko, N., Sastre, R., Gallardo, A., & Román, J. S. (1999). Self-curing membranes of chitosan/PAA IPNs obtained by radical polymerization: Preparation, characterization and interpolymer complexation. *Biomaterials, 20*(20), 1869-1878.

Rae, I. B., Gibb, S. W., & Lu, S. (2009). Biosorption of Hg from aqueous solutions by crab carapace. *Journal of Hazardous Materials, 164*(2-3), 1601-1604.

Rosa, S., Laranjeira, M. C. M., Riela, H. G., & Fávere, V. T. (2008). Cross-linked quaternary chitosan as an adsorbent for the removal of the reactive dye from aqueous solutions. *Journal of Hazardous Materials, 155*(1-2), 253-260.

Subramani, S. E., & Thinakaran, N. (2017). Isotherm, kinetic and thermodynamic studies on the adsorption behavior of textile dyes onto chitosan. *Process Safety and Environmental Protection, 106*, 1-10.

Szyguła, A., Guibal, E., Palacín, M. A., Ruiz, M., & Sastre, A. M. (2009). Removal of an anionic dye (acid blue 92) by coagulation–flocculation using chitosan. *Journal of Environmental Management, 90*(10), 2979-2986.

Vijayaraghavan, K., Palanivelu, K., & Velan, M. (2006). Biosorption of copper(II) and cobalt(II) from aqueous solutions by crab shell particles. *Bioresource Technology, 97*(12), 1411-1419.

REFEREES FOR THE PERTANIKA JOURNAL OF SCIENCE AND TECHNOLOGY

VOL. 26 (4) OCT. 2018

The Editorial Board of the Journal of Science and Technology wishes to thank the following for acting as referees for manuscripts published in this issue of JST.

A. Saravanakumar
(GMRIT, India)

Abdul Aziz Abdul Raman
(UM, Malaysia)

Achutha Kini
(MIT, India)

Aminuddin Ab Ghani
(USM, Malaysia)

Antony Selvadoss Thanamani
(NGM College, India)

Anusha Achuthan
(USM, Malaysia)

Arien Heryansyah
(UTM, Malaysia)

Arun Kumar Singh
(SMIT, India)

Arunodaya Raj Mishra
(ITM University, India)

Bujang Kim Huat
(UPM, Malaysia)

Che Zalina Zulkifli
(UPSI, Malaysia)

Chelli N Devi
(VIT University, India)

Cherry Bhargava
(IKGPTU, India)

Chua Han Bing
(Curtin Unviersity, Malaysia)

D.K. Singh
(IARI, Indonesia)

Divakar Yadav
(MMMUT, India)

Dzati Athiar Ramli
(USM, Malaysia)

Faridah Othman
(UM, Malaysia)

Gan Chin Kim
(UTeM, Malaysia)

Ghasem Najafpour
(NIT, Iran)

Hamidah Ibrahim
(UPM, Malaysia)

Hari Mohan Srivastava
(UViC, Canada)

Heng Chin Low
(USM, Malaysia)

Hima Bindu Maringanti
(NOU, India)

Isa Bala Muhammad
(FUTMinna, Nigeria)

Juraidah Ahmad
(UiTM, Malaysia)

K. Kannan
(SASTRA University, India)

Koh You Beng
(UM, Malaysia)

Lam Man Kee
(UTP, Malayia)

Law Ming Chiat
(Curtin University, Malaysia)

Lilly Suriani Affendey
(UPM, Malaysia)

Lim Chin Wai
(UPM, Malaysia)

M Iqbal Saripan
(UPM, Malaysia)

Manoj Kumar
(LPU, India)

Marcus Jopony
(UMS, Malaysia)

Masnida Hussin
(UPM, Malaysia)

Md Azlin Md Said
(USM, Malaysia)

Mohammad Abdul Mojid
(BAU, Malaysia)

Mohammed A. Alhanjouri
(IUGAZA, Palestine)

Mohd Rosli Hainin
(UTM, Malaysia)

Moola Mohan Reddy
(Curtin University, Malaysia)

Muhammad Ekhlasur Rahman
(Curtin University, Malaysia)

Narayanan Kumar
(Vels University, India)

Narongrit Sombatsompop
(KMUTT, Thailand)

Neeraj Kumar
(Thapar Unviersity, India)

Noor Atinah Ahmad
(USM, Malaysia)

Nor Ashidi Mat Isa
(USM, Malaysia)

AUN	*: American University of Nigeria*
BAU	*: Bangladesh Agricultural University*
CGC	*: Chandigarh Group of Colleges*
FUTMinna	*: Federal University of Technology, Minna*
GMRIT	*: GMR Institute of Technology*
IARI	*: Indian Agricultural Research Institute*
IIT	*: Indian Institute of Technology*
IKGPTU	*: IK Gujral Punjab Technical University*
ITM University	*: Institute of Technology & Management University*
IUGAZA	*: Islamic University of Gaza*
JUIT	*: Jaypee University of Information Technology*
KMUTT	*: King Mongkut's University of Technology Thonburi*
LPU	*: Lovely Professional University*
MIT	*: Manipal Institute of Technology*
MMMUT	*: Madan Mohan Malaviya University of Technology*
NGM College	*: Nallamuthu Gounder Mahalingam College*
NIT	*: Noshirvani University of Technology*
NIT Kurukshetra	*: National Institute of Technology, Kurukshetra*
NOU	*: North Orissa University*
PEC	*: Punjab Engineering College*
SASTRA University	*: Shanmugha Arts, Science, Technology & Research Academy University*
SMIT	*: Sikkim Manipal Institute of Technology*
UiTM	*: Universiti Teknologi MARA*
UM	*: Universiti Malaya*
UMP	*: Universiti Malaysia Pahang*
UMS	*: Universiti Malaysia Sabah*
UniKL	*: University of Kuala Lumpur*
UNIMAS	*: University Malaysia Sarawak*
UNITEN	*: Universiti Tenaga Nasional*
UPM	*: Universiti Putra Malaysia*
UPSI	*: Universiti Pendidikan Sultan Idris*
USM	*: Universiti Sains Malaysia*
UTeM	*: Universiti Teknikal Malaysia Melaka*
UTM	*: Universiti Teknologi Malaysia*
UTP	*: Universiti Teknologi Petronas*
UVic	*: University of Victoria*
VIT University	*: Vellore Institute of Technology University*

While every effort has been made to include a complete list of referees for the period stated above, however if any name(s) have been omitted unintentionally or spelt incorrectly, please notify the Chief Executive Editor, *Pertanika* Journals at executive_editor.pertanika@upm.my

Any inclusion or exclusion of name(s) on this page does not commit the *Pertanika* Editorial Office, nor the UPM Press or the University to provide any liability for whatsoever reason.

PERTANIKA JOURNALS

Pertanika Journals

Our goal is to bring high quality research to the widest possible audience

INSTRUCTIONS TO AUTHORS

(Manuscript Preparation & Submission Guide)

Revised: August 2018

Please read the Pertanika guidelines and follow these instructions carefully. Manuscripts not adhering to the instructions will be returned for revision without review. The Chief Executive Editor reserves the right to return manuscripts that are not prepared in accordance with these guidelines.

MANUSCRIPT PREPARATION

Manuscript Types

Pertanika accepts submission of mainly **four** types of manuscripts for peer-review.

1. **Regular article**

Regular articles are full-length original empirical investigations, consisting of introduction, materials and methods, results and discussion, conclusions. Original work must provide references and an explanation on research findings that contain new and significant findings.

Size: Generally, these are expected to be between 6 and 12 journal pages (excluding the abstract, references, tables and/or figures), a maximum of 80 references, and an abstract of 100–200 words.

2. **Review article**

These report critical evaluation of materials about current research that has already been published by organizing, integrating, and evaluating previously published materials. It summarizes the status of knowledge and outline future directions of research within the journal scope. Review articles should aim to provide systemic overviews, evaluations and interpretations of research in a given field. Re-analyses as meta-analysis and systemic reviews are encouraged. The manuscript title must start with "Review Article:".

Size: These articles do not have an expected page limit or maximum number of references, should include appropriate figures and/or tables, and an abstract of 100–200 words. Ideally, a review article should be of 7 to 8 printed pages.

3. **Short communications**

They are timely, peer-reviewed and brief. These are suitable for the publication of significant technical advances and may be used to:

(a) report new developments, significant advances and novel aspects of experimental and theoretical methods and techniques which are relevant for scientific investigations within the journal scope;

(b) report/discuss on significant matters of policy and perspective related to the science of the journal, including 'personal' commentary;

(c) disseminate information and data on topical events of significant scientific and/or social interest within the scope of the journal.

The manuscript title must start with "*Brief Communication*:".

Size: These are usually between 2 and 4 journal pages and have a maximum of three figures and/or tables, from 8 to 20 references, and an abstract length not exceeding 100 words. Information must be in short but complete form and it is not intended to publish preliminary results or to be a reduced version of Regular or Rapid Papers.

4. **Others**

Brief reports, case studies, comments, concept papers, Letters to the Editor, and replies on previously published articles may be considered.

PLEASE NOTE: NO EXCEPTIONS WILL BE MADE FOR PAGE LENGTH.

Language Accuracy

Pertanika **emphasizes** on the linguistic accuracy of every manuscript published. Articles must be in **English** and they must be competently written and argued in clear and concise grammatical English. Contributors are strongly advised to have the manuscript checked by a colleague with ample experience in writing English manuscripts or a competent English language editor.

Author(s) **must provide a certificate** confirming that their manuscripts have been adequately edited. A proof from a recognised editing service should be submitted together with the cover letter at the time of submitting a manuscript to Pertanika**. All editing costs must be borne by the author(s)**. This step, taken by authors before submission, will greatly facilitate reviewing, and thus publication if the content is acceptable.

Linguistically hopeless manuscripts will be rejected straightaway (e.g., when the language is so poor that one cannot be sure of what the authors really mean). This process, taken by authors before submission, will greatly facilitate reviewing, and thus publication if the content is acceptable.

MANUSCRIPT FORMAT

The paper should be submitted in one column format with at least 4cm margins and 1.5 line spacing throughout. Authors are advised to use Times New Roman 12-point font and *MS Word* format.

1. Manuscript Structure

Manuscripts in general should be organised in the following order:

Page 1: *Running title*

This page should **only** contain the running title of your paper. The running title is an abbreviated title used as the running head on every page of the manuscript. The running title should not exceed 60 characters, counting letters and spaces.

Page 2: *Author(s) and Corresponding author information.*

This page should contain the **full title** of your paper not exceeding 25 words, with name(s) of all the authors, institutions and corresponding author's name, institution and full address (Street address, telephone number (including extension), hand phone number, and e-mail address) for editorial correspondence. First and corresponding authors must be clearly indicated.

The names of the authors may be abbreviated following the international naming convention. e.g. Salleh, A.B.[1], Tan, S.G[2*]., and Sapuan, S.M[3].

Authors' addresses. Multiple authors with different addresses must indicate their respective addresses separately by superscript numbers:

George Swan[1] and Nayan Kanwal[2]
[1]Department of Biology, Faculty of Science, Duke University, Durham, North Carolina, USA.,
[2]Office of the Deputy Vice Chancellor (R&I), Universiti Putra Malaysia, Serdang, Malaysia.

A **list** of number of **black and white / colour figures and tables** should also be indicated on this page**.** Figures submitted in color will be printed in colour. See "*5. Figures & Photographs"* for details.

Page 3: *Abstract*

This page should **repeat** the **full title** of your paper with only the **Abstract** (the abstract should be less than 250 words for a Regular Paper and up to 100 words for a Short Communication), and **Keywords**.

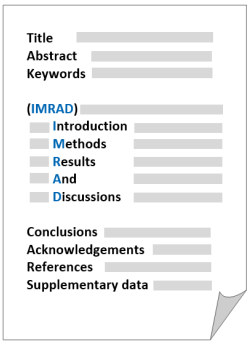
Keywords: Not more than eight keywords in alphabetical order must be provided to describe the contents of the manuscript.

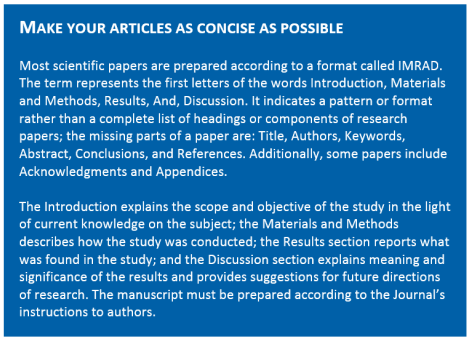
Page 4: *Introduction*

This page should begin with the **Introduction** of your article and followed by the rest of your paper.

2. *Text*

Regular Papers should be prepared with the headings *Introduction, Materials and Methods, Results and Discussion*, *Conclusions, Acknowledgements, References, and Supplementary data* (if avavailble) in this order.





3. *Equations and Formulae*

These must be set up clearly and should be typed double spaced. Numbers identifying equations should be in square brackets and placed on the right margin of the text.

4. *Tables*

All tables should be prepared in a form consistent with recent issues of Pertanika and should be numbered consecutively with Roman numerals. Explanatory material should be given in the table legends and footnotes. Each table should be prepared on a new page, embedded in the manuscript.

When a manuscript is submitted for publication, tables must also be submitted separately as data - .doc, .rtf, Excel or PowerPoint files- because tables submitted as image data cannot be edited for publication and are usually in low-resolution.

5. *Figures & Photographs*

Submit an **original** figure or photograph. Line drawings must be clear, with high black and white contrast. Each figure or photograph should be prepared on a new page, embedded in the manuscript for reviewing to keep the file of the manuscript under 5 MB. These should be numbered consecutively with Roman numerals.

Figures or photographs must also be submitted separately as TIFF, JPEG, or Excel files- because figures or photographs submitted in low-resolution embedded in the manuscript cannot be accepted for publication. For electronic figures, create your figures using applications that are capable of preparing high resolution TIFF files. In general, we require **300 dpi** or higher resolution for **coloured and half-tone artwork**, and **1200 dpi or higher** for **line drawings** are required.

Failure to comply with these specifications will require new figures and delay in publication.
NOTE: Illustrations may be produced in colour at no extra cost at the discretion of the Publisher; the author could be charged Malaysian Ringgit 50 for each colour page.

6. *References*

References begin on their own page and are listed in alphabetical order by the first author's last name. Only references cited within the text should be included. All references should be in 12-point font and double-spaced.

NOTE: When formatting your references, please follow the **APA reference style** (6th Edition). Ensure that the references are strictly in the journal's prescribed style, failing which your article will **not be accepted for peer-review**. You may refer to the *Publication Manual of the American Psychological Association* for further details (http://www.apastyle.org/).

PERTANIKA JOURNALS

7. General Guidelines

Abbreviations: Define alphabetically, other than abbreviations that can be used without definition. Words or phrases that are abbreviated in the introduction and following text should be written out in full the first time that they appear in the text, with each abbreviated form in parenthesis. Include the common name or scientific name, or both, of animal and plant materials.

Acknowledgements: Individuals and entities that have provided essential support such as research grants and fellowships and other sources of funding should be acknowledged. Contributions that do not involve researching (clerical assistance or personal acknowledgements) should **not** appear in acknowledgements.

Authors' Affiliation: The primary affiliation for each author should be the institution where the majority of their work was done. If an author has subsequently moved to another institution, the current address may also be stated in the footer.

Co-Authors: The commonly accepted guideline for authorship is that one must have substantially contributed to the development of the paper and share accountability for the results. Researchers should decide who will be an author and what order they will be listed depending upon their order of importance to the study. Other contributions should be cited in the manuscript's Acknowledgements.

Copyright Permissions: Authors should seek necessary permissions for quotations, artwork, boxes or tables taken from other publications or from other freely available sources on the Internet before submission to Pertanika. Acknowledgement must be given to the original source in the illustration legend, in a table footnote, or at the end of the quotation.

Footnotes: Current addresses of authors if different from heading may be inserted here.

Page Numbering: Every page of the manuscript, including the title page, references, tables, etc. should be numbered.

Spelling: The journal uses American or British spelling and authors may follow the latest edition of the Oxford Advanced Learner's Dictionary for British spellings.

SUBMISSION OF MANUSCRIPTS

Owing to the volume of manuscripts we receive, we must insist that all submissions be made electronically using the **online submission system ScholarOne™**, a web-based portal by Thomson Reuters. For more information, go to our web page and click "**Online Submission**".

Submission Checklist

1. **MANUSCRIPT**: Ensure your MS has followed the Pertanika style particularly the first four pages as explained earlier. The article should be written in a good academic style and provide an accurate and succinct description of the contents ensuring that grammar and spelling errors have been corrected before submission. It should also not exceed the suggested length.

   COVER LETTER: All submissions must be accompanied by a cover letter detailing what you are submitting. Papers are accepted for publication in the journal on the understanding that the article is **original** and the content has **not been published** either **in English** or **any other language(s)** or **submitted for publication elsewhere**. The letter should also briefly describe the research you are reporting, why it is important, and why you think the readers of the journal would be interested in it. The cover letter must also contain an acknowledgement that all authors have contributed significantly, and that all authors have approved the paper for release and are in agreement with its content.

   The cover letter of the paper should contain (i) the title; (ii) the full names of the authors; (iii) the addresses of the institutions at which the work was carried out together with (iv) the full postal and email address, plus telephone numbers and emails of all the authors. The current address of any author, if different from that where the work was carried out, should be supplied in a footnote.

   The above must be stated in the cover letter. Submission of your manuscript will not be accepted until a cover letter has been received.

2. **COPYRIGHT**: Authors publishing the Journal will be asked to sign a copyright form. In signing the form, it is assumed that authors have obtained permission to use any copyrighted or previously published material. All authors must read and agree to the conditions outlined in the form, and must sign the form or agree that the corresponding author can sign on their behalf. Articles cannot be published until a signed form *(original pen-to-paper signature)* has been received.

Please do **not** submit manuscripts to the editor-in-chief or to any other office directly. Any queries must be directed to the **Chief Executive Editor's** office via email to executive_editor.pertanika@upm.my.

Visit our Journal's website for more details at http://www.pertanika.upm.edu.my/home.php.

HARDCOPIES OF THE JOURNALS AND OFF PRINTS

Under the Journal's open access initiative, authors can choose to download free material (via PDF link) from any of the journal issues from Pertanika's website. Under "**Browse Journals**" you will see a link, "*Current Issues*" or "*Archives*". Here you will get access to all current and back-issues from 1978 onwards.

Pertanika Journal of Science & Technology
Vol. 26 (4) Oct. 2018

Contents

Pertanika Editorial Office, Journal Division
Office of the Deputy Vice Chancellor (R&I),
1st Floor, IDEA Tower II,
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan
Malaysia

http://www.pertanika.upm.edu.my/
E-mail: executive_editor.pertanika@upm.my
Tel: +603 8947 1622/1616

PENERBIT UPM PRESS

http://penerbit.upm.edu.my
E-mail : penerbit@upm.edu.my
Tel : +603 8946 8855/8854

ISSN 0128-7680